GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 13:08:48 ; Search time 57.7 Seconds
(without alignments)
311.854 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 825
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	825	100.0	162	22	AAG98518	Human olfactory re
2	825	100.0	162	22	AAG72109	Human olfactory re
3	825	100.0	162	22	AAG72421	Human OR-like poly
4	806	97.7	162	22	AAG98500	Pan troglodytes ol
5	630	76.4	329	22	AAG71992	Human olfactory re
6	547	66.3	319	22	AAG71531	Human olfactory re
7	546	66.2	314	22	AAU24633	Human olfactory re
8	546	66.2	314	22	AAG71500	Human olfactory re
9	546	66.2	314	22	AAG72474	Human OR-like poly
10	527	63.9	162	22	AAG98487	Papio hamadryas ol
11	495	60.0	324	22	AAG71529	Human olfactory re
12	494	59.9	309	22	AAG71525	Human olfactory re
13	494	59.9	325	22	AAU24632	Human olfactory re
14	477.5	57.9	310	22	AAG71502	Human olfactory re
15	471	57.1	314	22	AAG71522	Human olfactory re
16	440	53.3	314	22	AAG71945	Human olfactory re
17	436	52.8	315	22	AAG71520	Human olfactory re
18	416	50.4	311	22	AAG71532	Human olfactory re
19	416	50.4	316	22	AAU24535	Human olfactory re
20	414	50.2	312	22	AAG73043	Olfactory receptor
21	411	49.8	314	22	AAB20487	Human G protein co
22	408	49.5	312	22	AAG73045	Olfactory receptor
23	407	49.3	312	22	AAG73044	Olfactory receptor
24	407	49.3	314	22	AAB20486	Human G protein co
25	406	49.2	320	22	AAG71832	Human olfactory re
26	401.5	48.7	333	22	AAG72682	Murine OR-like pol
27	400	48.5	301	22	AAG66341	Partial NOV 2 prot
28	400	48.5	312	22	AAG71610	Human olfactory re
29	400	48.5	312	22	AAG72568	Human OR-like poly
30	400	48.5	312	22	AAG73038	Olfactory receptor
31	400	48.5	314	22	AAU24552	Human olfactory re
32	400	48.5	314	22	AAG66322	Human NOV 2 protei
33	400	48.5	328	22	AAG71644	Human olfactory re
34	400	48.5	343	22	AAG72665	Murine OR-like pol
35	400	48.5	359	22	AAU24579	Human olfactory re
36	398	48.2	314	22	AAU24763	Human olfactory re
37	396	48.0	316	22	AAG71964	Human olfactory re
38	396	48.0	324	22	AAU24697	Human olfactory re
39	394	47.8	318	22	AAG73040	Olfactory receptor
40	393	47.6	317	22	AAU24664	Human olfactory re
41	393	47.6	317	22	AAG71834	Human olfactory re
42	392	47.5	310	22	AAG71819	Human olfactory re
43	391	47.4	313	22	AAG71718	Human olfactory re
44	391	47.4	318	22	AAG73042	Olfactory receptor
45	389	47.2	292	21	AAY90875	Human G protein-co

ALIGNMENTS

RESULT 1
AAG98518
ID AAG98518 standard; Protein; 162 AA.
XX
AC AAG98518;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human olfactory receptor 1.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance.
XX
OS Homo sapiens.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB02017.
XX
PR 22-DEC-1999; 99US-0171746.
PR 21-DEC-2000; 2000US-0747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR N-PSDB; AAH84009.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful
PT for analysis odours e.g. in food processing and perfumery -
XX

PS Claim 3; Page 266; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 825; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.8e-93;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
Db 1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
Db 61 lfkiscngpsinaliififgafiqiptlmtiiisytrvlfdilkkksekgrskafstcga 120

Qy 121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
Db 121 hllsvslyygtlifmyvrpasglaedqdkvyslfytiiipll 162

RESULT 2
AAG72109
ID AAG72109 standard; Protein; 162 AA.
XX
AC AAG72109;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1790.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 1178; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 825; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.8e-93;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
Db 1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
Db 61 lfkiscngpsinaliififgafiqiptlmtiiisytrvlfdilkkksekgrskafstcga 120

Qy 121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
Db 121 hllsvslyygtlifmyvrpasglaedqdkvyslfytiiipll 162

RESULT 3
AAG72421
ID AAG72421 standard; Protein; 162 AA.
XX
AC AAG72421;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human OR-like polypeptide query sequence, SEQ ID NO: 2102.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Example 6; Page 1411-1412; 1857pp; English.
XX
CC The present sequence is a polypeptide encoded by one of 344 newly mined
CC human genes. It was used as a query sequence in a database search of
CC olfactory receptor (OR)-like sequences. The invention relates to isolated
CC polynucleotides encoding polypeptides involved in olfactory sensation.
CC The polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary scents
CC and the identification of the odour receptors used to detect these
CC primary scents. The methods also enable determination of secondary scents

```
CC   and the identification of combinations of odour receptors that are
CC   involved in detecting such secondary scents. This enables the
CC   construction of a scent representation (also called a scent fingerprint
CC   or scent profile), which may be used to re-create and edit scents.
CC   Libraries of olfactory receptors are useful for determining the
CC   interaction pattern of a composition with the receptors, and can be
CC   used for determining differences in the olfactory faculties of different
CC   individuals.
XX
SQ   Sequence   162 AA;


  Query Match             100.0%;  Score 825;  DB 22;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-93;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 lfkiscngpsinaliififgafiqiptlmtiiisytrvlfdilkkksekgrskafstcga 120

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||||||||||||||||||||||||||||||||||||
Db    121 hllsvslyygtlifmyvrpasglaedqdkvyslfytiiipll 162


RESULT   4
AAG98500
ID   AAG98500 standard; Protein; 162 AA.
XX
AC   AAG98500;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Pan troglodytes olfactory receptor 9.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance.
XX
OS   Pan troglodytes.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   N-PSDB; AAH83985.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 3; Page 226; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence   162 AA;


  Query Match             97.7%;  Score 806;  DB 22;  Length 162;
  Best Local Similarity   97.5%;  Pred. No. 6e-91;
  Matches  158;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          |||||||||||||:||||||||||||||||||:||||||||||||||||||||||||| |
Db     61 lfkiscngpsinalmififgafiqiptlmtiiisysrvlfdilkkksekgrskafstcsa 120

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||||||||||||||||| ||||||||||||||
Db    121 hllsvslyygtlifmyvrpasglaedpdkvyslfytiiipll 162


RESULT   5
AAG71992
ID   AAG71992 standard; Protein; 329 AA.
XX
AC   AAG71992;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1673.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 1092-1093; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   329 AA;
```

```
  Query Match           76.4%;  Score 630;  DB 22;  Length 329;
  Best Local Similarity   74.7%;  Pred. No. 6.4e-69;
  Matches  121;  Conservative   13;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||||||:||| |  | : ||| |||||  :|| || |||||| ||||||||||||
Db     124 vaicnpllypvvmsnslctqfigisyfigflhsaihvgllfrltfcrsniihyfyceilq 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||||||  |::| |:|||| |||| | ||:|:||: :|  ||||||||||||||||| |
Db     184 lfkisctnptvnillififsafiqvftfmtlivsysyilsailkkksekgrskafstcsa 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |||||||:|||| ||||   || | || |:||||||||||||
Db     244 hllsvslfygtlffmyvssrsgsaadqakmyslfytiiipll 285


RESULT   6
AAG71531
ID   AAG71531 standard; Protein; 319 AA.
XX
AC   AAG71531;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1212.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 722; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   319 AA;


  Query Match           66.3%;  Score 547;  DB 22;  Length 319;
  Best Local Similarity   64.8%;  Pred. No. 9.9e-59;
  Matches  105;  Conservative   21;  Mismatches   36;  Indels .  0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||||:|:| |   || :|:: |||| |:|  :| |||||  ||||:|||:|:
Db     124 vaickpllypvimtnglciwllvlsfiggflhalihegilfrltfcnsniihhfycdiip 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | ||||  |||| |::||    ||: |::|:::||| ||| ||||||  |   ||||||||
Db     184 llkisctdpsinflmlfilsgsiqvftiltvlvsytfvlftilkkksakdirkafstcga 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |||||||||| |:|||| |||  |:||| | |||||:||| |
Db     244 hllsvslyygpllfmyvhpaspqaddqdmveslfytviipfl 285


RESULT   7
AAU24633
ID   AAU24633 standard; Protein; 314 AA.
XX
AC   AAU24633;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human olfactory receptor AOLFR127.
XX
KW   Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW   food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS   Homo sapiens.
XX
PN   WO200168805-A2.
XX
PD   20-SEP-2001.
XX
PF   13-MAR-2001; 2001WO-US07771.
XX
PR   13-MAR-2000; 2000US-0188914.
PR   24-MAR-2000; 2000US-0192033.
PR   12-APR-2000; 2000US-0198474.
PR   24-APR-2000; 2000US-0199335.
PR   26-MAY-2000; 2000US-0207702.
PR   23-JUN-2000; 2000US-0213849.
PR   16-AUG-2000; 2000US-0226534.
PR   07-SEP-2000; 2000US-0230732.
PR   07-FEB-2001; 2001US-0266862.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zozulya S;
XX
DR   WPI; 2001-570867/64.
DR   N-PSDB; AAS42326.
XX
PT   Nucleic acids encoding human olfactory G protein-coupled receptors,
PT   useful for screening for compounds involved in olfactory sensation,
PT   where the compounds can be used in the food, pharmaceutical and
PT   cosmetic industries to customise odours -
XX
PS   Claim 60; Page 136; 319pp; English.
XX
CC   The invention relates to nucleic acids encoding human olfactory
CC   receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC   specifically recognise molecules, odourants, that elicit specific
CC   olfactory sensation. The human olfactory receptors and polynucleotides
CC   encoding them are useful for screening a library of chemical compounds
CC   for compounds that are involved in olfactory sensation. Modulators of
CC   their activity are useful for pharmacological and genetic modulation of
CC   olfactory signalling pathways. Therefore, they can be used in the food,
CC   pharmaceutical and cosmetic industries to customise odours and
CC   fragrances. The present sequence is a human olfactory receptor of the
CC   invention.
XX
```

SQ Sequence 314 AA;

Query Match 66.2%; Score 546; DB 22; Length 314;
Best Local Similarity 64.8%; Pred. No. 1.3e-58;
Matches 105; Conservative 23; Mismatches 34; Indels 0; Gaps 0;

```
Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          |||| |||||||:|:| |  :||: |:: |||| |:|  |: |||||  ||||:|||:|:
Db    129 vaickpllypvimnnslcirllafsflggflhalihevlifrltfcnsniihhfycdiip 188

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          || |||  |||| |::||    ||: |::|:: |||  || |||||| :|  ||||||||
Db    189 lfmisctdpsinflmvfilsgsiqvftivtvlnsytfalftilkkksvrgvrkafstcga 248

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||| |||||:|||| |:||| : |:||||||||
Db    249 hllsvslyygplifmylrpaspqaddqdmidsvfytiiipll 290
```

RESULT 8
AAG71500
ID AAG71500 standard; Protein; 314 AA.
XX
AC AAG71500;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1181.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 697-698; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 314 AA;

Query Match 66.2%; Score 546; DB 22; Length 314;
Best Local Similarity 64.8%; Pred. No. 1.3e-58;
Matches 105; Conservative 23; Mismatches 34; Indels 0; Gaps 0;

```
Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          |||| |||||||:|:| |  :||: |:: |||| |:|  |: |||||  ||||:|||:|:
Db    129 vaickpllypvimnnslcirllafsflggflhalihevlifrltfcnsniihhfycdiip 188

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          || |||  |||| |::||    ||: |::|:: |||  || |||||| :|  ||||||||
Db    189 lfmisctdpsinflmvfilsgsiqvftivtvlnsytfalftilkkksvrgvrkafstcga 248

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||| |||||:|||| |:||| : |:||||||||
Db    249 hllsvslyygplifmylrpaspqaddqdmidsvfytiiipll 290
```

RESULT 9
AAG72474
ID AAG72474 standard; Protein; 314 AA.
XX
AC AAG72474;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human OR-like polypeptide query sequence, SEQ ID NO: 2155.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Example 6; Page 1448-1449; 1857pp; English.
XX
CC The present sequence is a polypeptide encoded by one of 344 newly mined
CC human genes. It was used as a query sequence in a database search of
CC olfactory receptor (OR)-like sequences. The invention relates to isolated
CC polynucleotides encoding polypeptides involved in olfactory sensation.
CC The polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary scents
CC and the identification of the odour receptors used to detect these
CC primary scents. The methods also enable determination of secondary scents
CC and the identification of combinations of odour receptors that are
CC involved in detecting such secondary scents. This enables the
CC construction of a scent representation (also called a scent fingerprint
CC or scent profile), which may be used to re-create and edit scents.
CC Libraries of olfactory receptors are useful for determining the
CC interaction pattern of a composition with the receptors, and can be
CC used for determining differences in the olfactory faculties of different
CC individuals.
XX
SQ Sequence 314 AA;

```
  Query Match           66.2%;  Score 546;  DB 22;  Length 314;
  Best Local Similarity  64.8%;  Pred. No. 1.3e-58;
  Matches  105;  Conservative   23;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| |||||:|:| |  :|: |:: |||| |:|  |: |||||  ||||:|||:|:
Db     129 vaickpllypvimnnslcirllafsflggflhalihevlifrltfcnsniihhfycdiip 188

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           || |||  |||| |::||    ||: |::|:: |||  || |||||| :|  ||||||||
Db     189 lfmisctdpsinflmvfilsgsiqvftivtvlnsytfalftilkkksvrgvrkafstcga 248

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |||||||||| |||||:||||  |:||| : |:|||||||||
Db     249 hllsvslyygplifmylrpaspqaddqdmidsvfytiiipll 290


RESULT  10
AAG98487
ID   AAG98487 standard; Protein; 162 AA.
XX
AC   AAG98487;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Papio hamadryas olfactory receptor 12.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance.
XX
OS   Papio hamadryas.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   N-PSDB; AAH83971.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 3; Page 199; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence   162 AA;


  Query Match           63.9%;  Score 527;  DB 22;  Length 162;
  Best Local Similarity  62.3%;  Pred. No. 1.2e-56;
  Matches  101;  Conservative   22;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| |||||:|:| |  :|| :|:| |||| |:|  :| |||||  ||||:|||:|:
Db       1 vaickpllypvimtnglcirllvlsfvggflhalihegilfrltfcnsniihhfycdiip 60

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |||  |||| |::||    ||: |::|:::||  ||| |||||| ||  ||||||||
Db      61 lltisctdpsinflmlfilsgsiqvftiltvlvsyafvlftilkkksvkgirkafstcga 120

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           || || |||| |:|||| |||  |:||| |  :|||:||| |
Db     121 hlfsvclyygpllfmyvgpaspqaddqdmvecvfytviipfl 162


RESULT  11
AAG71529
ID   AAG71529 standard; Protein; 324 AA.
XX
AC   AAG71529;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1210.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 720-721; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   324 AA;


  Query Match           60.0%;  Score 495;  DB 22;  Length 324;
  Best Local Similarity  61.7%;  Pred. No. 2.5e-52;
  Matches  100;  Conservative   19;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           ||||  |||||:|:| |  ||| :|:: | || |:|   | |||||  ||| :|||:|:
Db     124 vaickallypvimtnglciqllvlsfiggllhaliheiflfrltfcnsniiqhfycdiip 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
```

```
          | ||||    || |::|||   ||: |: ||:|||| ||  ||| || ||  || |||||
Db    184 llkisctdsfinflmvfifadsiqvftigtilisytlvllliilknksvkgiqkavstcga 243

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||| |:|||| ||  |:||| : |||||:|:|||
Db    244 hllsvslyygplvfmyvgsaspqaddqdmmeslfytvivpll 285


RESULT 12
AAG71525
ID   AAG71525 standard; Protein; 309 AA.
XX
AC   AAG71525;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1206.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 717 and 684; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   309 AA;


  Query Match             59.9%;  Score 494;  DB 22;  Length 309;
  Best Local Similarity   59.9%;  Pred. No. 3.1e-52;
  Matches   97;  Conservative   24;  Mismatches   41;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||  |||||:|:|:|  ||| :|:: | || |:| :   |||||  ||| :|||:|:
Db    124 vaickallypvimtnelciqllvlsfiggllhaliheafsfrltfcnsniiqhfycdiip 183

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          | ||||   ||| |::|||   :|: |: ||:|||| :|| ||:||| ||  || |||||
Db    184 llkisctdssinflmvfifagsvqvftigtilisytiilftilekksikgirkavstcga 243

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||| | | |:  ||  |:||| : |||||:|:|||
Db    244 hllsvslyygpltfkylgsaspqaddqdmmeslfytvivpll 285


RESULT 13
AAU24632
ID   AAU24632 standard; Protein; 325 AA.
XX
AC   AAU24632;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human olfactory receptor AOLFR126.
XX
KW   Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW   food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS   Homo sapiens.
XX
PN   WO200168805-A2.
XX
PD   20-SEP-2001.
XX
PF   13-MAR-2001; 2001WO-US07771.
XX
PR   13-MAR-2000; 2000US-0188914.
PR   24-MAR-2000; 2000US-0192033.
PR   12-APR-2000; 2000US-0198474.
PR   24-APR-2000; 2000US-0199335.
PR   26-MAY-2000; 2000US-0207702.
PR   23-JUN-2000; 2000US-0213849.
PR   16-AUG-2000; 2000US-0226534.
PR   07-SEP-2000; 2000US-0230732.
PR   07-FEB-2001; 2001US-0266862.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zozulya S;
XX
DR   WPI; 2001-570867/64.
DR   N-PSDB; AAS42325.
XX
PT   Nucleic acids encoding human olfactory G protein-coupled receptors,
PT   useful for screening for compounds involved in olfactory sensation,
PT   where the compounds can be used in the food, pharmaceutical and
PT   cosmetic industries to customise odours -
XX
PS   Claim 60; Page 136; 319pp; English.
XX
CC   The invention relates to nucleic acids encoding human olfactory
CC   receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC   specifically recognise molecules, odourants, that elicit specific
CC   olfactory sensation. The human olfactory receptors and polynucleotides
CC   encoding them are useful for screening a library of chemical compounds
CC   for compounds that are involved in olfactory sensation. Modulators of
CC   their activity are useful for pharmacological and genetic modulation of
CC   olfactory signalling pathways. Therefore, they can be used in the food,
CC   pharmaceutical and cosmetic industries to customise odours and
CC   fragrances. The present sequence is a human olfactory receptor of the
CC   invention.
XX
SQ   Sequence   325 AA;


  Query Match             59.9%;  Score 494;  DB 22;  Length 325;
  Best Local Similarity   59.9%;  Pred. No. 3.3e-52;
  Matches   97;  Conservative   24;  Mismatches   41;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||  |||||:|:|:|  ||| :|:: | || |:| :   |||||  ||| :|||:|:
Db    140 vaickallypvimtnelciqllvlsfiggllhaliheafsfrltfcnsniiqhfycdiip 199
```

```
Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | ||||   ||| |::|||   :|: |: ||:|||| :|| ||:||| ||  || |||||
Db     200 llkisctdssinflmvfifagsvqvftigtilisytiilftilekksikgirkavstcga 259

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |||||||||| | | |:  ||  |:||| : |||||:|:|||
Db     260 hllsvslyygpltfkylgsaspqaddqdmmeslfytvivpll 301


RESULT 14
AAG71502
ID   AAG71502 standard; Protein; 310 AA.
XX
AC   AAG71502;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1183.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 698-699; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   310 AA;


  Query Match             57.9%;  Score 477.5;  DB 22;  Length 310;
  Best Local Similarity   59.3%;  Pred. No. 3.4e-50;
  Matches   96;  Conservative   21;  Mismatches   40;  Indels    5;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:| |  :||      | || |:|   | |||||  |:::|: ||:|:
Db     124 vaickpllypaimtnglcirllg-----gllhalihegflfrltfcnsnvvhhiycdiip 178

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | ||||   ||| |::|||   ||: |: | :|||| ||| |||||| ||  |||||||
Db     179 lskisctdssinflmvfifsgsiqvftigtglisytfvlftilkkksvkgirkafstcga 238

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||||||||:| | |||:  ||  |:|:| : |||||:|:|||
Db     239 hllsvslyhgpldfmymgsaspqaddedmmeslfytvivpll 280


RESULT 15
AAG71522
ID   AAG71522 standard; Protein; 314 AA.
XX
AC   AAG71522;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1203.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 715-716; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   314 AA;


  Query Match             57.1%;  Score 471;  DB 22;  Length 314;
  Best Local Similarity   58.0%;  Pred. No. 2.2e-49;
  Matches   94;  Conservative   22;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||||:|:| |   || :|:: | || |:| :   |||||  | ||:| |:|:
Db     123 vaickpllypvimtnglciwlllslflggllhaliheafsyrltfcnsntihhfxcdiip 182

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | || |   |||  ::|||   ||: |: |:::||| ||  |||||| ||  | ||||||
Db     183 llkifctdssinfpmvfifscsiqvftigtvlvsytfvlstilkkksvkgirkdfstcga 242
```

```
Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |:| |||||| | ||||  ||  |:||| : |||||:|:|||
Db     243 hilpvslyygplafmyvgsasqraddqdmmeslfytvivpll 284
```

Search completed: June 27, 2002, 14:02:57
Job time: 3249 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 13:08:53 ; Search time 24.46 Seconds
(without alignments)
161.772 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 825
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	334	40.5	314	3	US-08-988-876-7	Sequence 7, Appli
2	318	38.5	284	1	US-08-118-270-67	Sequence 67, Appl
3	318	38.5	284	5	PCT-US93-08528-67	Sequence 67, Appl
4	312.5	37.9	296	2	US-08-467-948A-2	Sequence 2, Appli
5	312.5	37.9	296	3	US-08-467-947A-2	Sequence 2, Appli
6	311	37.7	222	2	US-08-467-948A-27	Sequence 27, Appl
7	311	37.7	222	3	US-08-467-947A-27	Sequence 27, Appl
8	308.5	37.4	277	1	US-08-118-270-68	Sequence 68, Appl
9	308.5	37.4	277	5	PCT-US93-08528-68	Sequence 68, Appl
10	304	36.8	316	2	US-08-827-291A-2	Sequence 2, Appli
11	293	35.5	247	1	US-08-465-980-3	Sequence 3, Appli
12	293	35.5	247	2	US-09-053-303-3	Sequence 3, Appli
13	293	35.5	247	5	PCT-US95-07093-3	Sequence 3, Appli
14	291	35.3	274	1	US-08-118-270-69	Sequence 69, Appl
15	291	35.3	274	5	PCT-US93-08528-69	Sequence 69, Appl
16	279	33.8	321	4	US-08-748-506-10	Sequence 10, Appl
17	279	33.8	321	4	US-08-748-506-18	Sequence 18, Appl
18	277	33.6	321	4	US-08-748-506-19	Sequence 19, Appl
19	276	33.5	309	3	US-08-988-876-5	Sequence 5, Appli
20	276	33.5	321	4	US-08-748-506-11	Sequence 11, Appl
21	273	33.1	321	4	US-08-748-506-13	Sequence 13, Appl
22	272	33.0	286	1	US-08-118-270-65	Sequence 65, Appl
23	272	33.0	286	5	PCT-US93-08528-65	Sequence 65, Appl
24	269	32.6	284	1	US-08-118-270-61	Sequence 61, Appl
25	269	32.6	284	5	PCT-US93-08528-61	Sequence 61, Appl
26	267	32.4	321	4	US-08-748-506-12	Sequence 12, Appl
27	267	32.4	321	4	US-08-748-506-20	Sequence 20, Appl
28	267	32.4	327	4	US-08-748-506-22	Sequence 22, Appl
29	267	32.4	327	4	US-08-748-506-23	Sequence 23, Appl
30	263	31.9	327	4	US-08-748-506-24	Sequence 24, Appl
31	259	31.4	333	3	US-08-988-876-6	Sequence 6, Appli
32	252	30.5	327	4	US-08-748-506-14	Sequence 14, Appl
33	250	30.3	275	1	US-08-118-270-66	Sequence 66, Appl
34	250	30.3	275	5	PCT-US93-08528-66	Sequence 66, Appl
35	237	28.7	269	1	US-08-118-270-64	Sequence 64, Appl
36	237	28.7	269	5	PCT-US93-08528-64	Sequence 64, Appl
37	231	28.0	273	1	US-08-118-270-63	Sequence 63, Appl
38	231	28.0	273	5	PCT-US93-08528-63	Sequence 63, Appl
39	228.5	27.7	277	1	US-08-118-270-62	Sequence 62, Appl
40	228.5	27.7	277	5	PCT-US93-08528-62	Sequence 62, Appl
41	219	26.5	293	1	US-08-118-270-60	Sequence 60, Appl
42	219	26.5	293	5	PCT-US93-08528-60	Sequence 60, Appl
43	171	20.7	177	4	US-08-748-506-21	Sequence 21, Appl
44	169	20.5	320	1	US-08-465-980-2	Sequence 2, Appli
45	169	20.5	320	2	US-09-053-303-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-988-876-7
; Sequence 7, Application US/08988876
; Patent No. 6063596
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    TITLE OF INVENTION:  G PROTEIN COUPLED RECEPTORS ASSOCIATED
;    TITLE OF INVENTION:  WITH IMMUNE RESPONSE
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,876
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0441 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  314 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  32086
US-08-988-876-7
```

```
  Query Match           40.5%;  Score 334;  DB 3;  Length 314;
  Best Local Similarity  40.1%;  Pred. No. 1e-29;
  Matches   65;  Conservative   38;  Mismatches   59;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| |   |:::|:|:   | ::|  |: || ||  |:| :|:|::
Db     124 VAICFPLHYTAIMSPMLCLALVALSWVLTTFHAMLHTLLMARLCFCADNVIPHFFCDMSA 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:: :   :|  :||| |  | :    : |: || |::  ||| | ||  |||||||:
Db     184 LLKLAFSDTRVNEWVIFIMGGLILVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |||:|||:| :|:  ::  :  :| | :: ||:: |:|
Db     244 HLSVVSLFYGTVIGLYLCSSANSSTLKDTVMAMMYTVVTPML 285


RESULT   2
US-08-118-270-67
; Sequence 67, Application US/08118270
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  284 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-67


  Query Match           38.5%;  Score 318;  DB 1;  Length 284;
  Best Local Similarity  38.3%;  Pred. No. 5.7e-28;
  Matches   62;  Conservative   42;  Mismatches   56;  Indels    2;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| |:   |: :|:|:   | ::|  :: ||:||  ::| :::|::
Db      96 VAICFPLHYMSIMSPKIZVSLVVLSWVLTTFHAMLHTLIMARLSFCEDSVIPHYFCDMST 155

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:: :   :|  :||| |  | :    : |: || |::  ||| | ||  |||||||:
Db     156 LLKVACSDTHDNELAIFILGGPIVVLPFLLIIVSYARIVSSIFKVPSSQSIHKAFSTCGS 215

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |||:|||:| :|:  ::  :  :| | :: ||:: |:|
Db     216 HLSVVSLFYGTVIGLYLCPSANNSEVKETVMSIY--TMVPML 255


RESULT   3
PCT-US93-08528-67
; Sequence 67, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  284 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-67


  Query Match.          38.5%;  Score 318;  DB 5;  Length 284;
  Best Local Similarity  38.3%;  Pred. No. 5.7e-28;
  Matches   62;  Conservative   42;  Mismatches   56;  Indels    2;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| |:   |: :|:|:   | ::|  :: ||:||  ::| :::|::
Db      96 VAICFPLHYMSIMSPKIZVSLVVLSWVLTTFHAMLHTLIMARLSFCEDSVIPHYFCDMST 155

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:: :   :|  :||| |  | :    : |: || |::  ||| | ||  |||||||:
Db     156 LLKVACSDTHDNELAIFILGGPIVVLPFLLIIVSYARIVSSIFKVPSSQSIHKAFSTCGS 215

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |||:|||:| :|:  ::  :  :| | :: ||:: |:|
Db     216 HLSVVSLFYGTVIGLYLCPSANNSEVKETVMSIY--TMVPML 255


RESULT   4
US-08-467-948A-2
; Sequence 2, Application US/08467948A
; Patent No. 5998164
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR2
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,948A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04079
;      FILING DATE:  30-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1140003/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  296 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-467-948A-2


  Query Match             37.9%;  Score 312.5;  DB 2;  Length 296;
  Best Local Similarity   41.1%;  Pred. No. 2.5e-27;
  Matches   67;  Conservative   35;  Mismatches   60;  Indels    1;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           ||||:|| | ::|: |:   |   |:  | |  :||||||:||| ||    |::|:|||| 
Db     127 VAICHPLRYFIIMTWKVCITLGITSWTCGSLLAMVHVSLILRLPFCGPREINHFFCEILS 186

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           : :::|    :| ::||    || :  |  :::||: :|  ||: :| :|| |||||| : 
Db     187 VLRLACADTWLNQVVIFEACMFILVGPLCLVLVSYSHILGGILRIQSGEGRRKAFSTCSS 246

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTII-IPLL 162
           ||  | |::|: | ||: | |   |:| ||  |    :  |:|
Db     247 HLCVVGLFFGSAIVMYMAPKSRHPEEQQKVLFLILQFLSTPML 289


RESULT   5
US-08-467-947A-2
; Sequence 2, Application US/08467947A
; Patent No. 6090575
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR1
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,947A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04079
;      FILING DATE:  30-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1140002/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  296 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-467-947A-2


  Query Match             37.9%;  Score 312.5;  DB 3;  Length 296;
  Best Local Similarity   41.1%;  Pred. No. 2.5e-27;
  Matches   67;  Conservative   35;  Mismatches   60;  Indels    1;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           ||||:|| | ::|: |:   |   |:  | |  :||||||:||| ||    |::|:|||| 
Db     127 VAICHPLRYFIIMTWKVCITLGITSWTCGSLLAMVHVSLILRLPFCGPREINHFFCEILS 186

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           : :::|    :| ::||    || :  |  :::||: :|  ||: :| :|| |||||| : 
Db     187 VLRLACADTWLNQVVIFEACMFILVGPLCLVLVSYSHILGGILRIQSGEGRRKAFSTCSS 246

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTII-IPLL 162
           ||  | |::|: | ||: | |   |:| ||  |    :  |:|
Db     247 HLCVVGLFFGSAIVMYMAPKSRHPEEQQKVLFLILQFLSTPML 289


RESULT   6
US-08-467-948A-27
; Sequence 27, Application US/08467948A
; Patent No. 5998164
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR2
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
```

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,948A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-467-948A-27

Query Match 37.7%; Score 311; DB 2; Length 222;
Best Local Similarity 38.9%; Pred. No. 2.6e-27;
Matches 61; Conservative 36; Mismatches 60; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
Db 64 VAICHPLRYSAIMSWRVCSTMAVTSWIIGVLLSLIHLVLLLPLPFCVSQKVNHFFCEITA 123

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
Db 124 ILKLACADTHLNETMVLAGAVSVLVGPFSSIVVSYACILGAILKIQSEEGQRKAFSTCSS 183

Qy 121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTI 157
Db 184 HLCVVGLFYGTAIVMYVGPRHGSPKEQKKYLLLFHSL 220

RESULT 7
US-08-467-947A-27
; Sequence 27, Application US/08467947A
; Patent No. 6090575
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; TITLE OF INVENTION: Coupled Receptor GPR1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,947A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-467-947A-27

Query Match 37.7%; Score 311; DB 3; Length 222;
Best Local Similarity 38.9%; Pred. No. 2.6e-27;
Matches 61; Conservative 36; Mismatches 60; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
Db 64 VAICHPLRYSAIMSWRVCSTMAVTSWIIGVLLSLIHLVLLLPLPFCVSQKVNHFFCEITA 123

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
Db 124 ILKLACADTHLNETMVLAGAVSVLVGPFSSIVVSYACILGAILKIQSEEGQRKAFSTCSS 183

Qy 121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTI 157
Db 184 HLCVVGLFYGTAIVMYVGPRHGSPKEQKKYLLLFHSL 220

RESULT 8
US-08-118-270-68
; Sequence 68, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:

```
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  277 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-68


  Query Match             37.4%;  Score 308.5;  DB 1;  Length 277;
  Best Local Similarity   37.7%;  Pred. No. 6.4e-27;
  Matches   61;  Conservative   36;  Mismatches   60;  Indels    5;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| |  | |: : :::   | |:|  |: ||:||  ::| :|:|:|
Db      95 VAICFPLRYTTIMSTKFCASLVLLLWMLTMRHALLHTLLIARLSFCEDSVILHFFCDISA 154

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||:   :| |:|:| |  | |   : |::|| |: | |||  | :   | |||||:
Db     155 LLKLSCSDIYVNELMIYILGGLIIIIPFLLIVMSYVRIFFSILKFPSIQDIYKVFSTCGS 214

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |:|:|||:  :|: |:       :       |:: |::
Db     215 HLSVVTLFYGTIFGIYLCPSG-----NNSTVKEILTVVTPMI 251


RESULT  9
PCT-US93-08528-68
; Sequence 68, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  277 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-68


  Query Match             37.4%;  Score 308.5;  DB 5;  Length 277;
  Best Local Similarity   37.7%;  Pred. No. 6.4e-27;
  Matches   61;  Conservative   36;  Mismatches   60;  Indels    5;  Gaps    1;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| |  | |: : :::   | |:|  |: ||:||  ::| :|:|:|
Db      95 VAICFPLRYTTIMSTKFCASLVLLLWMLTMRHALLHTLLIARLSFCEDSVILHFFCDISA 154

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||:   :| |:|:| |  | |   : |::|| |: | |||  | :   | |||||:
Db     155 LLKLSCSDIYVNELMIYILGGLIIIIPFLLIVMSYVRIFFSILKFPSIQDIYKVFSTCGS 214

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |:|:|||:  :|: |:       :       |:: |::
Db     215 HLSVVTLFYGTIFGIYLCPSG-----NNSTVKEILTVVTPMI 251


RESULT  10
US-08-827-291A-2
; Sequence 2, Application US/08827291A
; Patent No. 5874243
;  GENERAL INFORMATION:
;    APPLICANT:  Macina, Roberto
;    APPLICANT:  Sathe, Ganesh
;    TITLE OF INVENTION:  NOVEL OLRCC15 RECEPTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,291A
;      FILING DATE:  28-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  316 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-827-291A-2


  Query Match             36.8%;  Score 304;  DB 2;  Length 316;
  Best Local Similarity   36.0%;  Pred. No. 2.4e-26;
```

```
  Matches   58;  Conservative   38;  Mismatches   65;  Indels    0;  Gaps    0;

Qy      2 AICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQL 61
          |||:|| |  :|| |:   : : |:::|    :::        ::|    | :|:||:  |
Db    125 AICHPLRYTNLMSPKICGLMTAFSWILGSTDGIIYAVATFSFSYCGSREIAHFFCELPSL 184

Qy     62 FKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
           :|||  ||   :|||     : :  :  || ||  |:  ::   | :|| |||:|| :|
Db    185 LILSCNDTSIFEKVIFICSIVMLVFPVAIIIASYAGVILAVIHMGSGEGRRKAFTTCSSH 244

Qy    122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |: | ::||  :|||::| |  :  |||: |:||||: |:|
Db    245 LMVVGMFYGAGLFMYIQPTSDRSPTQDKLVSVFYTILTPML 285


RESULT  11
US-08-465-980-3
; Sequence 3, Application US/08465980
; Patent No. 5756309
;  GENERAL INFORMATION:
;    APPLICANT:  Soppet, Daniel R.
;    APPLICANT:  Li, Yi
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR HPRAJ70
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,980
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-446
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-465-980-3


  Query Match             35.5%;  Score 293;  DB 1;  Length 247;
  Best Local Similarity   43.6%;  Pred. No. 3e-25;
  Matches   58;  Conservative   27;  Mismatches   48;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          |||| || |  :||  |   |:::|:|:   | ::|  |: || ||  |:| :|:|::
Db    113 VAICFPLHYTAIMSPMLCLALVALSWVLTTFHAMLHTLLMARLCFCADNVIPHFFCDMSA 172

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          | |:: :   :|  :||| |  | :   : |: || |::  |||  | ||  |||||||:
Db    173 LLKLAFSDTRVNEWVIFIMGGLILVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGS 232

Qy    121 HLLSVSLYYGTLI 133
          || |||:|||:|
Db    233 HLSVVSLFYGTVI 245


RESULT  12
US-09-053-303-3
; Sequence 3, Application US/09053303
; Patent No. 5948890
;  GENERAL INFORMATION:
;    APPLICANT:  Soppet, Daniel R.
;    APPLICANT:  Li, Yi
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR HPRAJ70
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/053,303
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/465,980
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-446
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-053-303-3


  Query Match             35.5%;  Score 293;  DB 2;  Length 247;
  Best Local Similarity   43.6%;  Pred. No. 3e-25;
  Matches   58;  Conservative   27;  Mismatches   48;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          |||| || |  :||  |   |:::|:|:   | ::|  |: || ||  |:| :|:|::
Db    113 VAICFPLHYTAIMSPMLCLALVALSWVLTTFHAMLHTLLMARLCFCADNVIPHFFCDMSA 172

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          | |:: :   :|  :||| |  | :   : |: || |::  |||  | ||  |||||||:
Db    173 LLKLAFSDTRVNEWVIFIMGGLILVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGS 232

Qy    121 HLLSVSLYYGTLI 133
          || |||:|||:|
Db    233 HLSVVSLFYGTVI 245


RESULT  13
PCT-US95-07093-3
; Sequence 3, Application PC/TUS9507093
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Soppet, Daniel R.
;    APPLICANT:  Li, Yi
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR HPRAJ70
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07093
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/465,980
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-446
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-07093-3


  Query Match             35.5%;  Score 293;  DB 5;  Length 247;
  Best Local Similarity   43.6%;  Pred. No. 3e-25;
  Matches   58;  Conservative   27;  Mismatches   48;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| |   |:::|:|:   | ::|  |: || ||  |:| :|:|::
Db     113 VAICFPLHYTAIMSPMLCLALVALSWVLTTFHAMLHTLLMARLCFCADNVIPHFFCDMSA 172

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:: :   :|  :||| |  | :   : |: || |::  |||  | ||  ||||||:
Db     173 LLKLAFSDTRVNEWVIFIMGGLILVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGS 232

Qy     121 HLLSVSLYYGTLI 133
           ||  |||:|||:|
Db     233 HLSVVSLFYGTVI 245


RESULT  14
US-08-118-270-69
; Sequence 69, Application US/08118270
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-69


  Query Match             35.3%;  Score 291;  DB 1;  Length 274;
  Best Local Similarity   41.0%;  Pred. No. 5.8e-25;
  Matches   66;  Conservative   36;  Mismatches   49;  Indels   10;  Gaps    5;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| ||   |  :|:|:   | ::|  :: ||:||    : :|:|:|.
Db      96 VAICFPLHYMSIMSPKLCVSLW-LSWVLTTFHAMLHTLIMARLSFCA--DLPHFFCDISP 152

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||:   :| |:||: |  | || :: ||:|| ||:  |||  | :|  | |||||:
Db     153 LLKLSCSDTHVNELVIFL-GLVIVIPFVL-IIVSYARVVASILKVPSVRGIHKIFSTCGS 210

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPL 161
           ||  |||:|||:| :|: |::  :  :: :     |:: ||
Db     211 HLSVVSLFYGTIIGLYLCPSANNSTVKETL-----TVVTPL 246


RESULT  15
PCT-US93-08528-69
; Sequence 69, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY-2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-69


  Query Match             35.3%;  Score 291;  DB 5;  Length 274;
  Best Local Similarity   41.0%;  Pred. No. 5.8e-25;
  Matches   66;  Conservative   36;  Mismatches   49;  Indels   10;  Gaps    5;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| ||   |  :|:|:   | ::|  :: ||:||    : :|:|:|
Db      96 VAICFPLHYMSIMSPKLCVSLW-LSWVLTTFHAMLHTLIMARLSFCA--DLPHFFCDISP 152

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||:   :| |:||: |  | || :: ||:|| ||:  |||  | :|  | |||||:
Db     153 LLKLSCSDTHVNELVIFL-GLVIVIPFVL-IIVSYARVVASILKVPSVRGIHKIFSTCGS 210

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPL 161
           ||  |||:|||:| :|: |::  :  :: :     |:: ||
Db     211 HLSVVSLFYGTIIGLYLCPSANNSTVKETL-----TVVTPL 246


Search completed: June 27, 2002, 14:03:42
Job time: 3289 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 13:12:38 ; Search time 42.96 Seconds
(without alignments)
362.348 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 825
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	385	46.7	304	2	S29709	olfactory receptor
2	382	46.3	157	2	S58020	probable olfactory
3	382	46.3	309	1	S51356	olfactory receptor
4	371	45.0	157	2	S58017	probable olfactory
5	362	43.9	157	2	S57998	probable olfactory
6	360	43.6	314	2	A37286	olfactory receptor
7	358	43.4	157	2	S58038	probable olfactory
8	356	43.2	315	2	JC5201	chemoreceptor TB56
9	345	41.8	314	2	S29707	olfactory receptor
10	344	41.7	161	2	A40745	odorant receptor (
11	343	41.6	157	2	S58014	probable olfactory
12	342.5	41.5	313	2	S20571	olfactory receptor
13	341	41.3	157	2	S58007	probable olfactory
14	340	41.2	305	2	S29711	olfactory factor O
15	339	41.1	157	2	S58000	probable olfactory
16	338	41.0	315	2	JC4658	olfactory receptor
17	337	40.8	216	2	I38480	olfactory receptor
18	337	40.8	314	2	H23701	olfactory receptor
19	336	40.7	312	2	I23701	olfactory receptor
20	335	40.6	157	2	S58018	probable olfactory
21	334	40.5	312	2	S29708	olfactory receptor
22	334	40.5	314	2	S20572	olfactory receptor
23	333	40.4	157	2	S58030	probable olfactory
24	332	40.2	216	2	I38476	olfactory receptor
25	331	40.1	225	2	I38478	olfactory receptor
26	329	39.9	310	2	E23701	olfactory receptor
27	328	39.8	216	2	I38474	olfactory receptor
28	326.5	39.6	158	2	S58016	probable olfactory
29	326	39.5	157	2	S58026	probable olfactory
30	324	39.3	157	2	S58024	probable olfactory
31	324	39.3	216	2	I38470	olfactory receptor
32	322.5	39.1	156	2	S58072	probable olfactory
33	320	38.8	157	2	S58034	probable olfactory
34	320	38.8	216	2	I38479	olfactory receptor
35	318	38.5	157	2	S58027	probable olfactory
36	316	38.3	157	2	S57999	probable olfactory
37	316	38.3	312	2	G23701	olfactory receptor
38	313	37.9	157	2	S58068	probable olfactory
39	311	37.7	216	2	I38477	olfactory receptor
40	311	37.7	222	2	D40745	odorant receptor (
41	310	37.6	311	2	JC5200	chemoreceptor TB33
42	310	37.6	313	2	B23701	olfactory receptor
43	310	37.6	318	2	JC5202	chemoreceptor TB64
44	309	37.5	157	2	S57996	probable olfactory
45	308.5	37.4	312	2	A48413	probable olfactory

ALIGNMENTS

RESULT 1
S29709
olfactory receptor OR14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 04-Sep-1998
C;Accession: S29709
R;Raming, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Br
Nature 361, 353-356, 1993
A;Title: Cloning and expression of odorant receptors.
A;Reference number: S29707; MUID:93149273
A;Accession: S29709
A;Molecule type: mRNA
A;Residues: 1-304 <RAM>
C;Superfamily: olfactory receptor OR14

Query Match 46.7%; Score 385; DB 2; Length 304;
Best Local Similarity 48.5%; Pred. No. 3.8e-31;
Matches 79; Conservative 32; Mismatches 50; Indels 2; Gaps 2;

```
Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||:| | || ::   ||  :||:||   : |   |: ||||  |::::| |:||
Db     117 VAICNPLMYTVTMSPQVCLLLLLGAYVMGFSEAMAHTGNLMNLTFCADNLVNHFMCDILP 176

Qy      61 LFKISCNGPSINALIIFIFGAF-IQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCG 119
           | ::|||   || |::||  |  | :| :::| |||  :|  ||:  | :|||||||||
Db     177 LLELSCNSTFINELVVFIVVAIDIAVP-IVSIFISYALILSSILRMHSTEGRSKAFSTCS 235

Qy     120 AHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           :||: | | :|:  |||::  | |  || || ||||||::|:|
Db     236 SHLIVVCLLFGSGAFMYLKLPSILPLDQGKVSSLFYTIVVPML 278
```

RESULT 2
S58020
probable olfactory receptor tpcr09 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58020
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific repe
A;Reference number: S57995
A;Accession: S58020
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89681; NID:g902351; PIDN:CAA61828.1; PID:g902352
C;Superfamily: olfactory receptor OR14

Query Match 46.3%; Score 382; DB 2; Length 157;

```
  Best Local Similarity   40.8%;  Pred. No. 4e-31;
  Matches   64;  Conservative   46;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
           || |||| |:||::| : |: :::::||:  | | | :: |:||: :|| :::|::| |
Db       1 ICRPLLYNVIMSSRLCSLLVLVAFILGFVSALAHTSAMMNLSFCKSHIISHYFCDVLPLL 60

Qy      63 KISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAHL 122
             :||:   :| |::|| | |   :   : : |||  :   ||  || ||||||:| || :||
Db      61 NLSCSNTHLNELLLFIIGGFNTLVPTLAVAISYVFIFCSILHIKSSKGRSKSFGTCSSHL 120

Qy     123 LSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
           ::| :::|:: ||| :|:|  : :|:|| |:||| :|
Db     121 MAVGIFFGSITFMYFKPSSSNSLEQEKVSSVFYTTVI 157


RESULT   3
S51356
olfactory receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S51356; S47014
R;Gat, U.; Nekrasova, E.; Lancet, D.; Natochin, M.
Eur. J. Biochem. 225, 1157-1168, 1994
A;Title: Olfactory receptor proteins. Expression, characterization and partial purificat
A;Reference number: S51356; MUID:95045546
A;Accession: S51356
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-309 <GAT>
A;Cross-references: EMBL:X80671; NID:g517365; PIDN:CAA56697.1; PID:g517366
R;Gat, U.; Nekrasova, E.; Lancet, D.; Natochin, M.
submitted to the EMBL Data Library, July 1994
A;Description: Olfactory receptor proteins: expression, characterization and partial pur
A;Reference number: S47014
A;Accession: S47014
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-309 <GA2>
A;Cross-references: EMBL:X80671; NID:g517365; PIDN:CAA56697.1; PID:g517366
C;Superfamily: olfactory receptor OR14


  Query Match             46.3%;  Score 382;  DB 1;  Length 309;
  Best Local Similarity   40.1%;  Pred. No. 7.8e-31;
  Matches   65;  Conservative   48;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| |||| |:||::: : |: ::: :|    :|| | :: |:||:  || :::|: |
Db     125 VAICRPLLYNVIMSSRICSLLVLVAFSLGLFSAVVHTSAMMNLSFCKSYIISHYFCDALP 184

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |::|:   :| |:||| |    :   : : |||  :   ||: :| :|||||| || :
Db     185 LLKLACSNTHLNELLIFIIGGLNTLVPTLAVAISYVFIFCSILRIRSSEGRSKAFGTCSS 244

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::| :::|:: |||::|:|  : :|:|| |:||| :||:|
Db     245 HLMAVGIFFGSITFMYLKPSSSNSLEQEKVSSVFYTTVIPML 286


RESULT   4
S58017
probable olfactory receptor tpcr25 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58017
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58017
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89672; NID:g902327; PIDN:CAA61819.1; PID:g902328
C;Superfamily: olfactory receptor OR14


  Query Match             45.0%;  Score 371;  DB 2;  Length 157;
  Best Local Similarity   44.9%;  Pred. No. 5e-30;
  Matches   71;  Conservative   41;  Mismatches   44;  Indels    2;  Gaps    2;

Qy       3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
           ||:|||| |::|||    |:   |:|| :   ||   : |: ||:|::|::::|::| |
Db       1 ICSPLLYSVIISNKACFSLILGVYIIGLVCASVHTGCMFRVQFCKFDLINHYFCDLLPLL 60

Qy      63 KISCNGPSINALIIFIFGAF-IQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
           |:||:   :| |:|   ||| | :|:| ||: ||  ::  ||  :| :|||||||| :|
Db      61 KLSCSSIYVNKLLILCVGAFNILVPSL-TILCSYIFIIASILHIRSTEGRSKAFSTCSSH 119

Qy     122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
           :|:| :::|:  |||::|:|  : || || |:|||||:
Db     120 MLAVVIFFGSAAFMYLQPSSISSMDQGKVSSVFYTIIV 157


RESULT   5
S57998
probable olfactory receptor tpcr79 - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S57998
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific repe
A;Reference number: S57995
A;Accession: S57998
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89665; NID:g902198; PIDN:CAA61812.1; PID:g902199
C;Superfamily: olfactory receptor OR14


  Query Match             43.9%;  Score 362;  DB 2;  Length 157;
  Best Local Similarity   41.4%;  Pred. No. 4e-29;
  Matches   65;  Conservative   43;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
           ||:|||| |:|::    |:|  |:|  :   ||   : |: ||:|::|::::|::| |
Db       1 ICSPLLYNVIMAHQACFSLISGVYIIALVCAFVHTGCMFRVHFCKFDVINHYFCDLLSLL 60

Qy      63 KISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAHL 122
           |:||:   :| |:| ||     :  :||: ||  :|  ||  :| :|||||||| :|:
Db      61 KLSCSNTHVNELVILIFSTVNILAPSLTILASYVFILSSILHIRSTEGRSKAFSTCSSHI 120

Qy     123 LSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
            :|::::|:  |||::|:|  : || || |:|||||:
Db     121 SAVAVFFGSAAFMYLQPSSVSSMDQGKVSSVFYTIIV 157


RESULT   6
A37286
olfactory receptor I15 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
C;Accession: A37286
R;Buck, L.; Axel, R.
Cell 65, 175-187, 1991
A;Title: A novel multigene family may encode odorant receptors: a molecular basis for
A;Reference number: A23701; MUID:91191556
A;Accession: A37286
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-314 <BUC>
A;Cross-references: GB:M64392; NID:g205845; PIDN:AAA41755.1; PID:g205846
C;Superfamily: olfactory receptor OR14
```

C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 43.6%; Score 360; DB 2; Length 314;
Best Local Similarity 42.0%; Pred. No. 1.3e-28;
Matches 68; Conservative 39; Mismatches 55; Indels 0; Gaps 0;

```
Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| ||   |: :|:|:   | ::|  |: ||:|| |:| :|:|:|
Db     124 VAICFPLHYMSIMSPKLCVSLVVLSWVLTTFHAMLHTLLMARLSFCADNMIPHFFCDISP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||:   :| |:||: |  : :   : ||:|| ||:  ||| | :|  | |||||:
Db     184 LLKLSCSDTHVNELVIFVMGGLVIVIPFVLIIVSYARVVASILKVPSVRGIHKIFSTCGS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |||:|||:| :|: |::  :  :: | :: ||:: |:|
Db     244 HLSVVSLFYGTIIGLYLCPSANNSTVKETVMAMMYTVVTPML 285
```

RESULT 7
S58038
probable olfactory receptor tpcr21 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58038
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58038
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89704; NID:g902724; PIDN:CAA61851.1; PID:g902725
C;Superfamily: olfactory receptor OR14

Query Match 43.4%; Score 358; DB 2; Length 157;
Best Local Similarity 40.1%; Pred. No. 1e-28;
Matches 63; Conservative 45; Mismatches 49; Indels 0; Gaps 0;

```
Qy       3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
           |||||||   || ::   |:|  |:       |::  :||: ||: ::|::::|: | ||
Db       1 ICNPLLYNATMSFQIYCSLISGVYIFAVFCASVNLGFMLRIQFCKLDVINHYFCDFLPLF 60

Qy      63 KISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAHL 122
           :::|:   || ::|  ||       |:|:|||  ::  ||: :: :|||||||| :|:
Db      61 QLACSNTYINEMLILFFGILNIFVPLLTVIISYIFIIASILRIRTTEGRSKAFSTCSSHI 120

Qy     123 LSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
           |:|::::|:| |||::|:|  : || || |:||| ::
Db     121 LAVAIFFGSLAFMYLQPSSVSSMDQGKVSSVFYTTVV 157
```

RESULT 8
JC5201
chemoreceptor TB567 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-Aug-1999
C;Accession: JC5201; PC4303
R;Thomas, M.B.; Haines, S.L.; Akeson, R.A.
Gene 178, 1-5, 1996
A;Title: Chemoreceptors expressed in taste, olfactory and male reproductive tissues.
A;Reference number: JC5200; MUID:97080538
A;Accession: JC5201
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-315 <THO1>
A;Cross-references: GB:U50948; NID:g1256390; PIDN:AAC52910.1; PID:g1256391
A;Accession: PC4303
A;Status: preliminary
A;Molecule type: protein
A;Residues: 144-151;270-277 <THO2>
A;Experimental source: taste bud
C;Comment: This protein is coupled to a GTP-binding protein-mediated signal transduct
C;Genetics:
A;Gene: tb567
C;Superfamily: olfactory receptor OR14
C;Keywords: olfaction; taste bud; transmembrane protein
F;26-49/Domain: transmembrane #status predicted <TM1>
F;58-79/Domain: transmembrane #status predicted <TM2>
F;101-120/Domain: transmembrane #status predicted <TM3>
F;140-164/Domain: transmembrane #status predicted <TM4>
F;197-219/Domain: transmembrane #status predicted <TM5>
F;238-260/Domain: transmembrane #status predicted <TM6>
F;273-293/Domain: transmembrane #status predicted <TM7>

Query Match 43.2%; Score 356; DB 2; Length 315;
Best Local Similarity 38.9%; Pred. No. 3.3e-28;
Matches 63; Conservative 43; Mismatches 56; Indels 0; Gaps 0;

```
Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || ||::|| ::   |:::||::| ::  ||      |::|    |::|:|:::
Db     124 VAICKPLRYPIIMSRQVCLLLVALSYLMGSINSSVHTGFTFSLSYCNSKNINHFFCDVVP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           :  :||:   || ::: :|  |    |:: || ||  ::  |||  |  || | |||| :
Db     184 IISLSCSNTDINIMLLIVFVGFNLTFTVLVIIFSYIYIMAAILKMSSTAGRKKTFSTCAS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           || :|:::|||| :||::| |  :|:  ||  :|| |:||:|
Db     244 HLTAVTIFYGTLSYMYLQPHSDNSEENMKVAVVFYGIVIPML 285
```

RESULT 9
S29707
olfactory receptor OR5 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C;Accession: S29707; B37286
R;Raming, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Br
Nature 361, 353-356, 1993
A;Title: Cloning and expression of odorant receptors.
A;Reference number: S29707; MUID:93149273
A;Accession: S29707
A;Molecule type: mRNA
A;Residues: 1-314 <RAM>
R;Buck, L.; Axel, R.
Cell 65, 175-187, 1991
A;Title: A novel multigene family may encode odorant receptors: a molecular basis for
A;Reference number: A23701; MUID:91191556
A;Accession: B37286
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 193-236 <BUC>
A;Cross-references: GB:M64375; NID:g205811; PIDN:AAA41738.1; PID:g205812
C;Superfamily: olfactory receptor OR14
C;Keywords: G protein-coupled receptor; membrane protein

Query Match 41.8%; Score 345; DB 2; Length 314;
Best Local Similarity 41.4%; Pred. No. 4.1e-27;
Matches 67; Conservative 39; Mismatches 56; Indels 0; Gaps 0;

```
Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| || |  :|| ||   |: :|:|:   | ::|  |: ||:|| |:| :|:|::
Db     124 VAICFPLHYMSIMSPKLCVSLVVLSWVLTTFHAMLHTLLMARLSFCEDNVIPHFFCDMSA 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |::|:   :| ::|||  :   :     ||:|| |:: |||  | :|  ||||:||:
Db     184 LLKLACSDTRVNEVVIFIVVSLFLVLPFALIIMSYVRIVSSILKVPSSQGIYKAFSSCGS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |||:|||:| :|: |:|  :  :: | || ||:: |:|
```

```
Db    244 HLSVVSLFYGTVIPLYLCPSSNNSTVKETVMSLMYTLVTPML 285
```

RESULT 10
A40745
odorant receptor (clone M50) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: A40745
R;Ressler, K.J.; Sullivan, S.L.; Buck, L.B.
Cell 73, 597-609, 1993
A;Title: A zonal organization of odorant receptor gene expression in the olfactory epith
A;Reference number: A40745; MUID:93258822
A;Accession: A40745
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-161 <RES>
A;Cross-references: GB:L14568; NID:g293757; PIDN:AAA39852.1; PID:g293758
A;Experimental source: olfactory epithelium
A;Note: sequence extracted from NCBI backbone (NCBIP:131747)
C;Superfamily: olfactory receptor OR14
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 41.7%; Score 344; DB 2; Length 161;
Best Local Similarity 40.3%; Pred. No. 2.7e-27;
Matches 64; Conservative 37; Mismatches 58; Indels 0; Gaps 0;

```
Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          |||| || | |:|| |:   |:  ||:::||   : |   ::||:|| |||::: |:|
Db      3 VAICKPLTYKVIMSPKICCLLIFSSYLMGFASAMAHTGCMIRLSFCDSNIINHYMCDIFP 62

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          |  :||:   :| |:  :      |   : |:|||  :||:|:   | || |||  |||:
Db     63 LLPLSCSSTYVNELMSSVVVGSAIILCCLIILISYAMILFNIIHMSSGKGWSKALGTCGS 122

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
          |:::|||:||: :  ||:|:|     | | :|:|||:::
Db    123 HIITVSLFYGSGLLAYVKPSSAKTVGQGKFFSVFYTLLV 161
```

RESULT 11
S58014
probable olfactory receptor tpcr120 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58014
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58014
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89669; NID:g902321; PIDN:CAA61816.1; PID:g902322
C;Superfamily: olfactory receptor OR14

Query Match 41.6%; Score 343; DB 2; Length 157;
Best Local Similarity 40.1%; Pred. No. 3.3e-27;
Matches 63; Conservative 41; Mismatches 53; Indels 0; Gaps 0;

```
Qy      3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
          ||:|||| |:||      |    |::| |   :|   :|||  |: |:|::::|::  |
Db      1 ICSPLLYNVIMSYHHCFWLTVGVYILGILGSTIHTGFMLRLFLCKTNVINHYFCDLFPLL 60

Qy     63 KISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAHL 122
            :||:   || |:: :  ||  :   :||: ||  ::  ||: :| :|||||||||| :|:
Db     61 GLSCSSTYINELLVLVLSAFNILTPALTILASYIFIIASILRIRSTEGRSKAFSTCSSHI 120

Qy    123 LSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
          |:|: ::|:  |||::|:|  : || || |:||||::
```

```
Db    121 LAVAGFFGSAAFMYLQPSSVSSMDQGKVSSVFYTIVV 157
```

RESULT 12
S20571
olfactory receptor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 26-Aug-1999
C;Accession: S20571
R;Parmentier, M.; Libert, F.; Schurmans, S.; Schiffmann, S.; Lefort, A.; Eggerickx, D
Nature 355, 453-455, 1992
A;Title: Expression of members of the putative olfactory receptor gene family in mamm
A;Reference number: S20571; MUID:92131132
A;Accession: S20571
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-313 <PAR>
A;Cross-references: EMBL:X64996; NID:g890; PIDN:CAA46129.1; PID:g891
C;Superfamily: olfactory receptor OR14
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 41.5%; Score 342.5; DB 2; Length 313;
Best Local Similarity 41.4%; Pred. No. 7.3e-27;
Matches 67; Conservative 36; Mismatches 58; Indels 1; Gaps 1;

```
Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          |||| || |  :|| ||   || :|:|:   | ::|  |: || ||  | | :|:|::
Db    124 VAICFPLHYTTIMSPKLCFSLLVLSWVLTMFHAVLHTLLMARLCFCA-NTIPHFFCDMSA 182

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          | |::|:   :| |:|| |  | :   : || || |::  |||  |  |  | |||||:
Db    183 LLKLACSDTQVNELVIFIMGGLILVIPFLLIITSYARIVSSILKVPSAIGICKVFSTCGS 242

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          ||  |||:|||:| :|: |::  :  :: : :: ||:: |:|
Db    243 HLSVVSLFYGTVIGLYLCPSANNSTVKETIMAMMYTVVTPML 284
```

RESULT 13
S58007
probable olfactory receptor tpcr85 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58007
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific repe
A;Reference number: S57995
A;Accession: S58007
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89675; NID:g902333; PIDN:CAA61822.1; PID:g902334
C;Superfamily: olfactory receptor OR14

Query Match 41.3%; Score 341; DB 2; Length 157;
Best Local Similarity 43.0%; Pred. No. 5.2e-27;
Matches 68; Conservative 37; Mismatches 51; Indels 2; Gaps 2;

```
Qy      3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
          ||||||| | || ::   ||   | :||   : | : :: :|||  |:::: |:|| |
Db      1 ICNPLLYMVTMSPQVCFLLLLGVYGMGFAGAMAHTACMMGVTFCANNLVNHYMCDILPLL 60

Qy     63 KISCNGPSINALIIFIFGAF-IQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
          : :|     :| |:::|:  :  | :|| :|| |||  :|  |     | :|||||||| :|
Db     61 ECACTSTYVNELVVFVVVSIDIGVPT-VTIFISYALILSSIFHIDSTEGRSKAFSTCSSH 119

Qy    122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
          :::|||::|:  |||::| | || :| || ||||| ::
Db    120 IIAVSLFFGSGAFMYLKPFSLLAMNQGKVSSLFYTTVV 157
```

RESULT 14
S29711
olfactory factor OR37 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 04-Sep-1998
C;Accession: S29711
R;Raming, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Breer
Nature 361, 353-356, 1993
A;Title: Cloning and expression of odorant receptors.
A;Reference number: S29707; MUID:93149273
A;Accession: S29711
A;Molecule type: mRNA
A;Residues: 1-305 <RAM>
C;Superfamily: olfactory receptor OR14

Query Match 41.2%; Score 340; DB 2; Length 305;
Best Local Similarity 41.4%; Pred. No. 1.3e-26;
Matches 70; Conservative 34; Mismatches 57; Indels 8; Gaps 3;

```
Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||| | |:|| ::   : | |:  | :: :|  || :|| ||  |:|::| ||:|
Db    112 VAICNPLRYSVIMSKEVYVSMASASWFSGGINSVVQTSLAMRLPFCGNNVINHFTCEVLA 171

Qy     61 LFKISCNGPSINALIIFIFG-AFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCG 119
          : |::|   |:| : : |   ||: :| |: |  ||  :|: ||:  |  || ||||||
Db    172 VLKLACADISLNIVTMVISNMAFLVLP-LLLIFFSYVLILYTILRMNSASGRRKAFSTCS 230

Qy    120 AHLLSVSLYYGTLIFMYVRPAS------GLAEDQDKVYSLFYTIIIPLL 162
          |||  | ::|||:  || :| |         :  ||: |||| :: |:|
Db    231 AHLTVVVIFYGTIFSMYAKPKSQDLTGKDKFQTSDKIISLFYGVVTPML 279
```

RESULT 15
S58000
probable olfactory receptor tpcr71 - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 03-Nov-2000
C;Accession: S58000
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58000
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89663; NID:g902194; PIDN:CAA61810.1; PID:g902195
C;Superfamily: olfactory receptor OR14

Query Match 41.1%; Score 339; DB 2; Length 157;
Best Local Similarity 43.2%; Pred. No. 8.2e-27;
Matches 67; Conservative 34; Mismatches 54; Indels 0; Gaps 0;

```
Qy      3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
          ||:|| | |:|| ::       |:  | |  |||| |:||| ||  : |::|:|||| :
Db      1 ICHPLHYSVIMSWRVCTVQAVTSWACGSLLALVHVILILRLPFCGPHEINHFFCEILSVL 60

Qy     63 KISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAHL 122
          |::|    :| ::||     | :  |  :::||:|:|| ||: :| :|| |||||| :||
Db     61 KLACADTRLNQVVIFAASVSILVGPLCLVLVSYSRILFAILRIQSGEGRRKAFSTCSSHL 120

Qy    123 LSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTI 157
            | |::|: | ||: | |   |:| |: ||||::
Db    121 CVVGLFFGSAIVMYMAPKSNHPEEQQKILSLFYSL 155
```

Search completed: June 27, 2002, 14:04:46
Job time: 3128 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:03:48 ; Search time 23.96 Seconds
(without alignments)
261.793 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 825
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	414	50.2	312	1	OLF4_CHICK	P37070 gallus gall
2	408	49.5	312	1	OLF6_CHICK	P37072 gallus gall
3	407	49.3	312	1	OLF5_CHICK	P37071 gallus gall
4	394	47.8	318	1	OLF1_CHICK	P37067 gallus gall
5	391	47.4	318	1	OLF3_CHICK	P37069 gallus gall
6	388	47.0	312	1	OLF2_CHICK	P37068 gallus gall
7	379	45.9	311	1	OLF1_CANFA	Q95154 canis famil
8	379	45.9	314	1	O5I1_HUMAN	Q13606 homo sapien
9	374	45.3	311	1	OLF2_CANFA	Q95155 canis famil
10	366	44.4	311	1	O8B8_HUMAN	Q15620 homo sapien
11	366	44.4	311	1	O8D2_HUMAN	Q9gzm6 homo sapien
12	360	43.6	314	1	OLF1_RAT	P23274 rattus norv
13	347	42.1	314	1	O5F1_HUMAN	O95221 homo sapien
14	346	41.9	315	1	O3A3_HUMAN	P47888 homo sapien
15	346	41.9	321	1	O5V1_HUMAN	Q9ugf6 homo sapien
16	344	41.7	161	1	OL7A_MOUSE	P34985 mus musculu
17	342.5	41.5	313	1	OLFD_CANFA	P30955 canis famil
18	338	41.0	315	1	O3A1_HUMAN	P47881 homo sapien
19	337	40.8	314	1	OLF9_RAT	P23272 rattus norv
20	337	40.8	323	1	O1E2_HUMAN	P47887 homo sapien
21	336	40.7	312	1	OLF4_RAT	P23273 rattus norv
22	334	40.5	314	1	O1E1_HUMAN	P30953 homo sapien
23	331	40.1	312	1	O1F1_HUMAN	O43749 homo sapien
24	329	39.9	310	1	OLF0_RAT	P23269 rattus norv
25	328	39.8	159	1	OLF8_CHICK	Q98913 gallus gall
26	328	39.8	216	1	O1E5_HUMAN	Q9um60 homo sapien
27	324	39.3	216	1	O5D4_HUMAN	Q9up62 homo sapien
28	323	39.2	313	1	O1C1_HUMAN	Q15619 homo sapien
29	320	38.8	216	1	OL1I_HUMAN	P47889 homo sapien
30	320	38.8	313	1	O2B6_HUMAN	P58173 homo sapien
31	320	38.8	315	1	O3A2_HUMAN	P47893 homo sapien
32	320	38.8	316	1	O2G1_HUMAN	Q9h205 homo sapien
33	318	38.5	319	1	O2S2_HUMAN	Q9nqn1 homo sapien
34	317	38.4	313	1	O1G1_HUMAN	P47890 homo sapien
35	316	38.3	312	1	OLF8_RAT	P23271 rattus norv
36	316	38.3	313	1	OLF5_RAT	P23266 rattus norv
37	314	38.1	312	1	OAC1_HUMAN	Q96kk4 homo sapien
38	313	37.9	357	1	O2B2_HUMAN	Q9gzk3 homo sapien
39	311	37.7	216	1	OL1F_HUMAN	P47886 homo sapien
40	311	37.7	222	1	OL13_MOUSE	P34984 mus musculu
41	311	37.7	314	1	O1Q1_HUMAN	Q15612 homo sapien
42	309	37.5	355	1	O1I1_HUMAN	O60431 homo sapien
43	308.5	37.4	312	1	O1D2_HUMAN	P34982 homo sapien
44	307	37.2	313	1	O2B3_HUMAN	O76000 homo sapien
45	305	37.0	348	1	O3A4_HUMAN	P47883 homo sapien

ALIGNMENTS

```
RESULT   1
OLF4_CHICK
ID   OLF4_CHICK      STANDARD;      PRT;   312 AA.
AC   P37070;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor-like protein COR4.
GN   COR4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=96317247; PubMed=8734500;
RA   Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT   "Olfaction in birds: differential embryonic expression of nine
RT   putative odorant receptor genes in the avian olfactory system.";
RL   Mech. Dev. 55:65-77(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z79593; CAB01854.1; -.
DR   GCRDb; GCR_1095; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Olfaction; Multigene family.
FT   DOMAIN        1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    164       4 (POTENTIAL).
FT   DOMAIN      165    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    260       6 (POTENTIAL).
FT   DOMAIN      261    271       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM    272    292       7 (POTENTIAL).
FT   DOMAIN      293    312       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     97    179       BY SIMILARITY.
SQ   SEQUENCE   312 AA;  35093 MW;  BA893613AC929BC1 CRC64;


  Query Match             50.2%;  Score 414;  DB 1;  Length 312;
  Best Local Similarity   45.1%;  Pred. No. 1.2e-31;
  Matches   73;  Conservative   47;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:  :  :|:   | : ||: ||| | ||:|:|| |::::|:|:
Db     124 VAICKPLLYPSIMTKAVCWRLVKGLYSLAFLNSLVHTSGLLKLSFCSSNVVNHFFCDNSP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:|| :  ::| |::|||||:   : :::||:||| ::  ::: :|: |: |||||| :
Db     184 LFQISSSSTTLNELLVFIFGSLFAMSSIITILISYVFIILTVVRIRSKDGKYKAFSTCTS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::|||::||:||||:||    : | ||: |||||::||:|
Db     244 HLMAVSLFHGTVIFMYLRPVKLFSLDTDKIASLFYTVVIPML 285


RESULT   2
OLF6_CHICK
ID   OLF6_CHICK     STANDARD;      PRT;   312 AA.
AC   P37072; Q98912;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor-like protein COR6.
GN   COR6.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=96317247; PubMed=8734500;
RA   Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT   "Olfaction in birds: differential embryonic expression of nine
RT   putative odorant receptor genes in the avian olfactory system.";
RL   Mech. Dev. 55:65-77(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z79590; CAB01851.1; -.
DR   GCRDb; GCR_1092; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Olfaction; Multigene family.
FT   DOMAIN        1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    164       4 (POTENTIAL).
FT   DOMAIN      165    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    260       6 (POTENTIAL).
FT   DOMAIN      261    271       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    292       7 (POTENTIAL).
FT   DOMAIN      293    312       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     97    179       BY SIMILARITY.
SQ   SEQUENCE   312 AA;  35179 MW;  56D164BF99D59EA9 CRC64;


  Query Match             49.5%;  Score 408;  DB 1;  Length 312;
  Best Local Similarity   43.5%;  Pred. No. 4.3e-31;
  Matches   70;  Conservative   51;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       2 AICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQL 61
           ||| ||||| :|:  :  :|:   | : ||: ||| | ||:|:|| |::::|:|:   |
Db     125 AICKPLLYPAIMTKAVCWRLVKGLYSLAFLNFLVHTSGLLKLSFCSSNVVNHFFCDNSPL 184

Qy      62 FKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
           |:|| :  ::| |::||||:   : :::||:|||  ::  ::: :|:: : ||||||:|
Db     185 FQISSSSTALNELLVFIFGSLFVMSSIITILISYVFIILTVVRIRSKERKYKAFSTCTSH 244

Qy     122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |::|||::||::||| :||:  : |:||: |||||::||:|
Db     245 LMAVSLFHGTIVFMYFQPANNFSLDKDKIMSLFYTVVIPML 285


RESULT   3
OLF5_CHICK
ID   OLF5_CHICK     STANDARD;      PRT;   312 AA.
AC   P37071;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor-like protein COR5.
GN   COR5.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=96317247; PubMed=8734500;
RA   Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT   "Olfaction in birds: differential embryonic expression of nine
RT   putative odorant receptor genes in the avian olfactory system.";
RL   Mech. Dev. 55:65-77(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z79589; CAB01850.1; -.
DR   GCRDb; GCR_1093; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
```

```
KW   Olfaction; Multigene family.
FT   DOMAIN        1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    164       4 (POTENTIAL).
FT   DOMAIN      165    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    260       6 (POTENTIAL).
FT   DOMAIN      261    271       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    292       7 (POTENTIAL).
FT   DOMAIN      293    312       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     97    179       BY SIMILARITY.
SQ   SEQUENCE   312 AA;  35256 MW;  FBC1BF1DC4542F0B CRC64;


  Query Match             49.3%;  Score 407;  DB 1;  Length 312;
  Best Local Similarity   44.4%;  Pred. No. 5.3e-31;
  Matches   72;  Conservative   47;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:  :  :|:   | : ||: ||| | ||:|:||  |::::|:|:
Db     124 VAICKPLLYPAIMTKAVCWRLVEGLYSLAFLNSLVHTSGLLKLSFCSSNVVNHFFCDNSP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:|| :  ::| |::||||::  : :::|  |||  ::  ::: :|: |: ||||||
Db     184 LFQISSSSTTLNELLVFIFGSWFAMSSIITTPISYVFIILTVVRIRSKDGKYKAFSTCTS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::|||::||:||||:||    : | ||: |||||::||:|
Db     244 HLMAVSLFHGTVIFMYLRPVKLFSLDTDKIASLFYTVVIPML 285


RESULT   4
OLF1_CHICK
ID   OLF1_CHICK     STANDARD;      PRT;   318 AA.
AC   P37067;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor-like protein COR1.
GN   COR1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=96317247; PubMed=8734500;
RA   Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT   "Olfaction in birds: differential embryonic expression of nine
RT   putative odorant receptor genes in the avian olfactory system.";
RL   Mech. Dev. 55:65-77(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z79584; CAB01845.1; -.
DR   GCRDb; GCR_1100; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Olfaction; Multigene family.
FT   DOMAIN        1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    164       4 (POTENTIAL).
FT   DOMAIN      165    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    260       6 (POTENTIAL).
FT   DOMAIN      261    271       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    292       7 (POTENTIAL).
FT   DOMAIN      293    318       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     97    179       BY SIMILARITY.
SQ   SEQUENCE   318 AA;  35668 MW;  C145276D16DE370D CRC64;


  Query Match             47.8%;  Score 394;  DB 1;  Length 318;
  Best Local Similarity   43.8%;  Pred. No. 8.8e-30;
  Matches   71;  Conservative   46;  Mismatches   45;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:  :  :|:   | : ||: ||| | ||:|:||  |::::|:|:|
Db     124 VAICKPLLYPSIMTKAVCWRLVESLYFLAFLNSLVHTSGLLKLSFCYSNVVNHFFCDISP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:|| :  :|: |:: | |:   : ::: |:|||  ::  ::  :|: |: ||||||
Db     184 LFQISSSSIAISELLVIISGSLFVMSSIIIILISYVFIILTVVMIRSKDGKYKAFSTCTS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::|||::||:||||:||    : | ||: |||||::||:|
Db     244 HLMAVSLFHGTVIFMYLRPVKLFSLDTDKIASLFYTVVIPML 285


RESULT   5
OLF3_CHICK
ID   OLF3_CHICK     STANDARD;      PRT;   318 AA.
AC   P37069;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor-like protein COR3.
GN   COR3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=96317247; PubMed=8734500;
RA   Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT   "Olfaction in birds: differential embryonic expression of nine
RT   putative odorant receptor genes in the avian olfactory system.";
RL   Mech. Dev. 55:65-77(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z79585; CAB01846.1; -.
DR   GCRDb; GCR_1097; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Olfaction; Multigene family.
FT   DOMAIN       1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27    49       1 (POTENTIAL).
FT   DOMAIN       50    57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58    79       2 (POTENTIAL).
FT   DOMAIN       80    100      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     101   120      3 (POTENTIAL).
FT   DOMAIN       121   139      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     140   164      4 (POTENTIAL).
FT   DOMAIN       165   205      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     206   226      5 (POTENTIAL).
FT   DOMAIN       227   239      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     240   260      6 (POTENTIAL).
FT   DOMAIN       261   271      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     272   292      7 (POTENTIAL).
FT   DOMAIN       293   318      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       5     5      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID      97   179      BY SIMILARITY.
SQ   SEQUENCE   318 AA;  35757 MW;  D50C72EBA043A962 CRC64;

  Query Match             47.4%;  Score 391;  DB 1;  Length 318;
  Best Local Similarity   43.2%;  Pred. No. 1.7e-29;
  Matches   70;  Conservative   46;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:  :  :|:   | : ||: |||   ||:|:|| |::::|:|:|
Db     124 VAICKPLLYPAIMTKAVCWRLVESLYFLAFLNSLVHTCGLLKLSFCYSNVVNHFFCDISP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:|| :  :|: |:: | |:   : ::: |:||| ::  ::  :|: |: |||||| :
Db     184 LFQISSSSIAISELLVIISGSLFVMSSIIIILISYVFIILTVVMIRSKDGKYKAFSTCTS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::|||::||:||||:||    : | ||: |||||::||:|
Db     244 HLMAVSLFHGTVIFMYLRPVKLFSLDTDKIASLFYTVVIPML 285


RESULT   6
OLF2_CHICK
ID   OLF2_CHICK     STANDARD;      PRT;   312 AA.
AC   P37068;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor-like protein COR2.
GN   COR2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=96317247; PubMed=8734500;
RA   Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT   "Olfaction in birds: differential embryonic expression of nine
RT   putative odorant receptor genes in the avian olfactory system.";
RL   Mech. Dev. 55:65-77(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z79588; CAB01849.1; -.
DR   GCRDb; GCR_1099; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Olfaction; Multigene family.
FT   DOMAIN       1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27    49       1 (POTENTIAL).
FT   DOMAIN       50    57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58    79       2 (POTENTIAL).
FT   DOMAIN       80    100      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     101   120      3 (POTENTIAL).
FT   DOMAIN       121   139      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     140   164      4 (POTENTIAL).
FT   DOMAIN       165   205      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     206   226      5 (POTENTIAL).
FT   DOMAIN       227   239      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     240   260      6 (POTENTIAL).
FT   DOMAIN       261   271      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     272   292      7 (POTENTIAL).
FT   DOMAIN       293   312      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       5     5      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID      97   179      BY SIMILARITY.
SQ   SEQUENCE   312 AA;  35255 MW;  3A4362DF2722770C CRC64;

  Query Match             47.0%;  Score 388;  DB 1;  Length 312;
  Best Local Similarity   43.2%;  Pred. No. 3.1e-29;
  Matches   70;  Conservative   46;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:  :  :|:   | : ||: ||| | ||:|:||  |::::|:|:
Db     124 VAICKPLLYPSIMTKAVCWRLVKGLYSLAFLNSLVHTSGLLKLSFCSSNVVNHFFCDNRP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:|| :  ::| |:: | |:   : :::||:|||  ::  ::  :|: |: ||||||  :
Db     184 LFQISSSSTTLNELLVIISGSLFVMSSIITILISYVFIILTVVMIRSKDGKYKAFSTCTS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::|||::||:||||:|     : | ||: |||||::||:|
Db     244 HLMAVSLFHGTVIFMYLRSVKLFSLDTDKIASLFYTVVIPML 285


RESULT   7
OLF1_CANFA
ID   OLF1_CANFA     STANDARD;      PRT;   311 AA.
AC   Q95154;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Olfactory receptor-like protein OLF1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=97008103; PubMed=8855279;
RA   Issel-Tarver L., Rine J.;
RT   "Organization and expression of canine olfactory receptor genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:10897-10902(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT OR SPERM CELL RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U53679; AAB37239.1; -.
DR   GCRDb; GCR_1192; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     24       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     25     48       1 (POTENTIAL).
FT   DOMAIN       49     56       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     57     78       2 (POTENTIAL).
FT   DOMAIN       79     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    100    119       3 (POTENTIAL).
FT   DOMAIN      120    138       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    139    157       4 (POTENTIAL).
FT   DOMAIN      158    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    218       5 (POTENTIAL).
FT   DOMAIN      219    235       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    236    259       6 (POTENTIAL).
FT   DOMAIN      260    271       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    291       7 (POTENTIAL).
FT   DOMAIN      292    311       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   311 AA;  35209 MW;  05A2132474F543A1 CRC64;

  Query Match             45.9%;  Score 379;  DB 1;  Length 311;
  Best Local Similarity   43.8%;  Pred. No. 2.1e-28;
  Matches   71;  Conservative   36;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||||| |:|| :  |: :||: | :  ||| |    | :|  |:|::|:|::
Db     123 VAICNPLLYTVVMSRGICVWLIVLSYIGGNMSSLVHTSFAFILKYCDKNVINHFFCDLPP 182

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||   |:|  ::  :|: ::|   : |:|||  :|  :|: :|  || | |||| :
Db     183 LLKLSCTDTSVNEWLLSTYGSSVEIFCFIVIVISYYFILRSVLRIRSSSGRKKTFSTCAS 242

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           || ||::| |||:|:| ||      : ||: |:|||||||:|
Db     243 HLTSVAIYQGTLLFIYSRPTYLYTPNTDKIISVFYTIIIPVL 284


RESULT   8
O5I1_HUMAN
ID   O5I1_HUMAN     STANDARD;      PRT;   314 AA.
AC   Q13606;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 5I1 (Olfactory receptor-like protein OLF1).
GN   OR5I1 OR OLF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Issel-Tarver L., Rine J.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U56420; AAB01214.1; -.
DR   GCRDb; GCR_1925; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     29     51       1 (POTENTIAL).
FT   DOMAIN       52     59       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     60     81       2 (POTENTIAL).
FT   DOMAIN       82    102       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    103    122       3 (POTENTIAL).
FT   DOMAIN      123    141       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    142    166       4 (POTENTIAL).
FT   DOMAIN      167    207       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    208    228       5 (POTENTIAL).
FT   DOMAIN      229    241       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    242    262       6 (POTENTIAL).
FT   DOMAIN      263    273       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    274    294       7 (POTENTIAL).
FT   DOMAIN      295    314       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    191       BY SIMILARITY.
FT   CARBOHYD      7      7       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   314 AA;  36048 MW;  334BE9BF3D9A4D63 CRC64;

  Query Match             45.9%;  Score 379;  DB 1;  Length 314;
  Best Local Similarity   44.4%;  Pred. No. 2.2e-28;
  Matches   72;  Conservative   37;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||||| |:||  :  :|: :||: | :  ||| |    | :|  |:|::|:|::
Db     126 VAICNPLLYTVVMSRGICMRLIVLSYLGGNMSSLVHTSFAFILKYCDKNVINHFFCDLPP 185

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||   :||  ::  :|: ::|   : |||||  :|  :|| :|  || | |||| :
Db     186 LLKLSCTDTTINEWLLSTYGSSVEIICFIIIIISYFFILLSVLKIRSFSGRKKTFSTCAS 245

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           || ||::| |||:|:| ||:   : : ||: |:|||| ||:|
Db     246 HLTSVTIYQGTLLFIYSRPSYLYSPNTDKIISVFYTIFIPVL 287


RESULT   9
OLF2_CANFA
ID   OLF2_CANFA     STANDARD;      PRT;   311 AA.
AC   Q95155;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Olfactory receptor-like protein OLF2.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97008103; PubMed=8855279;
RA   Issel-Tarver L., Rine J.;
RT   "Organization and expression of canine olfactory receptor genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:10897-10902(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT OR SPERM CELL RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U53680; AAB37240.1; -.
DR   GCRDb; GCR_1191; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    158       4 (POTENTIAL).
FT   DOMAIN      159    196       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    197    219       5 (POTENTIAL).
FT   DOMAIN      220    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    260       6 (POTENTIAL).
FT   DOMAIN      261    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    311       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   311 AA;  35177 MW;  C57D24B3B48FAC20 CRC64;


  Query Match             45.3%;  Score 374;  DB 1;  Length 311;
  Best Local Similarity   41.6%;  Pred. No. 6.3e-28;
  Matches   67;  Conservative   43;  Mismatches   51;  Indels    0;  Gaps    0;

Qy       2 AICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQL 61
           || :|||| | ||::: : |::  |::| :   |:  |  || ||  |:|::|:|::  |
Db     125 AISHPLLYTVSMSSRVCSLLMAGVYLVGIMDASVNTILTFRLCFCESNVINHFFCDVPPL 184

Qy      62 FKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
             :||:   :| |:||    ||:: ||  : :||  ::  : |  | :|| |||||| :|
Db     185 LLLSCSDTQVNELVIFTIFGFIELITLSGLFVSYCYIILAVRKINSAEGRFKAFSTCTSH 244

Qy     122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           | :|::: ||::||| ||:|  : ||||: ||||:::||:|
Db     245 LTAVAIFQGTMLFMYFRPSSSYSLDQDKIISLFYSLVIPML 285


RESULT  10
O8B8_HUMAN
ID   O8B8_HUMAN     STANDARD;      PRT;   311 AA.
AC   Q15620;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 8B8 (Olfactory receptor TPCR85).
GN   OR8B8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 126-282 FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X89675; CAA61822.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    158       4 (POTENTIAL).
FT   DOMAIN      159    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    219       5 (POTENTIAL).
FT   DOMAIN      220    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    259       6 (POTENTIAL).
FT   DOMAIN      260    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    311       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     97    189       BY SIMILARITY.
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    204    204       G -> S (IN REF. 1).
SQ   SEQUENCE   311 AA;  34482 MW;  9D8AAC4DA179ABE6 CRC64;


  Query Match             44.4%;  Score 366;  DB 1;  Length 311;
  Best Local Similarity   44.8%;  Pred. No. 3.5e-27;
  Matches   73;  Conservative   37;  Mismatches   51;  Indels    2;  Gaps    2;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||||| | || ::   ||   | :||   : | : :: :|||  |:::::: |:||
Db     124 VAICNPLLYMVTMSPQVCFLLLLGVYGMGFAGAMAHTACMMGVTFCANNLVNHYMCDILP 183

Qy      61 LFKISCNGPSINALIIF-IFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCG 119
           | : :|    :| |::| : |  | :|| :|| |||  :|  |    | :|||||||||
Db     184 LLECACTSTYVNELVVFVVVGIDIGVPT-VTIFISYALILSSIFHIDSTEGRSKAFSTCS 242

Qy     120 AHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           :|:::|||::|:  |||::| | || :| || ||||| ::|:|
Db     243 SHIIAVSLFFGSGAFMYLKPFSLLAMNQGKVSSLFYTTVVPML 285


RESULT  11
O8D2_HUMAN
```

```
ID   O8D2_HUMAN     STANDARD;      PRT;   311 AA.
AC   Q9GZM6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 8D2 (Olfactory receptor-like protein JCG2).
GN   OR8D2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tongue;
RA   Gaudin J.-C., Chobert J.-M., Haertle T.;
RT   "Molecular cloning of olfactory receptor-like mRNAs expressed in human
RT   tongue.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE TONGUE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF162668; AAG43386.1; -.
DR   EMBL; AF162669; AAG43387.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    158       4 (POTENTIAL).
FT   DOMAIN      159    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    219       5 (POTENTIAL).
FT   DOMAIN      220    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    259       6 (POTENTIAL).
FT   DOMAIN      260    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    311       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     97    189       BY SIMILARITY.
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   311 AA;  34857 MW;  ECE606FBBC962A82 CRC64;


  Query Match           44.4%;  Score 366;  DB 1;  Length 311;
  Best Local Similarity  39.5%;  Pred. No. 3.5e-27;
  Matches   64;  Conservative   45;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
             |||| |||| ::||::: : :::: | :||   || : :  |:||| : : :::|:||
Db     124 VAICRPLLYNIVMSHRVCSIMMAVVYSLGFLWATVHTTRMSVLSFCRSHTVSHYFCDILP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |  :||:   || :::|| |    : | : ::|||  :   ||   | :|:|||| || :
Db     184 LLTLSCSSTHINEILLFIIGGVNTLATTLAVLISYAFIFSSILGIHSTEGQSKAFGTCSS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
```

```
           |||:| :::|:: ||| :| |    :::|| |:||  |||:|
Db     244 HLLAVGIFFGSITFMYFKPPSSTTMEKEKVSSVFYITIIPML 285


RESULT  12
OLF1_RAT
ID   OLF1_RAT       STANDARD;      PRT;   314 AA.
AC   P23274;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Olfactory receptor-like protein I15.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91191556; PubMed=1840504;
RA   Buck L., Axel R.;
RT   "A novel multigene family may encode odorant receptors: a molecular
RT   basis for odor recognition.";
RL   Cell 65:175-187(1991).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: OLFACTORY EPITHELIUM.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M64392; AAA41755.1; -.
DR   PIR; A37286; A37286.
DR   GCRDb; GCR_0152; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    158       4 (POTENTIAL).
FT   DOMAIN      159    196       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM .  197    219       5 (POTENTIAL).
FT   DOMAIN      220    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    260       6 (POTENTIAL).
FT   DOMAIN      261    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    314       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     97    189       BY SIMILARITY.
SQ   SEQUENCE   314 AA;  35563 MW;  ED257FABB9A10CEB CRC64;


  Query Match           43.6%;  Score 360;  DB 1;  Length 314;
  Best Local Similarity  42.0%;  Pred. No. 1.3e-26;
  Matches   68;  Conservative   39;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
             |||| || |  :|| ||   |: :|:|:   | ::|  |: ||:||  |:| :|:|:|
Db     124 VAICFPLHYMSIMSPKLCVSLVVLSWVLTTFHAMLHTLLMARLSFCADNMIPHFFCDISP 183
```

```
Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |:||:   :| |:||: |  : :   : ||:|| ||:  ||| | :|  | |||||:
Db     184 LLKLSCSDTHVNELVIFVMGGLVIVIPFVLIIVSYARVVASILKVPSVRGIHKIFSTCGS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  |||:|||:| :|: |::  :  :: | :: ||:: |:|
Db     244 HLSVVSLFYGTIIGLYLCPSANNSTVKETVMAMMYTVVTPML 285


RESULT  13
O5F1_HUMAN
ID   O5F1_HUMAN     STANDARD;      PRT;   314 AA.
AC   O95221;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 5F1 (Olfactory receptor 11-10) (OR11-10).
GN   OR5F1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 68-283 FROM N.A.
RX   MEDLINE=99005533; PubMed=9787077;
RA   Buettner J.A., Glusman G., Ben-Arie N., Ramos P., Lancet D.,
RA   Evans G.A.;
RT   "Organization and evolution of olfactory receptor genes on human
RT   chromosome 11.";
RL   Genomics 53:56-68(1998).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF065863; AAC70017.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN       1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26    49       1 (POTENTIAL).
FT   DOMAIN       50    57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58    79       2 (POTENTIAL).
FT   DOMAIN       80    100      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     101   120      3 (POTENTIAL).
FT   DOMAIN       121   139      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     140   158      4 (POTENTIAL).
FT   DOMAIN       159   195      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     196   219      5 (POTENTIAL).
FT   DOMAIN       220   236      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     237   259      6 (POTENTIAL).
FT   DOMAIN       260   272      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     273   292      7 (POTENTIAL).
FT   DOMAIN       293   314      CYTOPLASMIC (POTENTIAL).
FT   DISULFID      97   189      BY SIMILARITY.
FT   CARBOHYD       5     5      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   314 AA;  35131 MW;  C14A9A35CD76E3A3 CRC64;


  Query Match           42.1%;  Score 347;  DB 1;  Length 314;
  Best Local Similarity 41.0%;  Pred. No. 2.1e-25;

  Matches   66;  Conservative   39;  Mismatches   56;  Indels    0;  Gaps    0;

Qy       2 AICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQL 61
           ||| |||| ::|| :  :: : ::  | |: :|: | :  |:|| |:||:|:|:   |
Db     125 AICRPLLYSLIMSRTVYLKMAAGAFAAGLLNFMVNTSHVSSLSFCDSNVIHHFFCDSPPL 184

Qy      62 FKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
           ||:||:   :   | |       : ||: |: ||: ||| |    | :|| :||||| :|
Db     185 FKLSCSDTILKESISSILAGVNIVGTLLVILSSYSYVLFSIFSMHSGEGRHRAFSTCASH 244

Qy     122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           | :: |:| | |: |:|||:|  : :|||| |:|||::||:|
Db     245 LTAIILFYATCIYTYLRPSSSYSLNQDKVASVFYTVVIPML 285


RESULT  14
O3A3_HUMAN
ID   O3A3_HUMAN     STANDARD;      PRT;   315 AA.
AC   P47888; Q9UBE7; Q9P1Q4;
DT   01-FEB-1996 (Rel. 33, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 3A3 (Olfactory receptor 17-201) (OR17-201).
GN   OR3A3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ferraz C., Vidal S., Brun C., Demaille J.G.;
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20139433; PubMed=10673334;
RA   Glusman G., Sosinsky A., Ben-Asher E., Avidan N., Sonkin D., Bahar A.,
RA   Rosenthal A., Clifton S., Roe B., Ferraz C., Demaille J.G., Lancet D.;
RT   "Sequence, structure, and evolution of a complete human olfactory
RT   receptor gene cluster.";
RL   Genomics 63:227-245(2000).
RN   [3]
RP   SEQUENCE OF 71-286 FROM N.A.
RX   MEDLINE=94272458; PubMed=8004088;
RA   Ben-Arie N., Lancet D., Taylor C., Khen M., Walker N.,
RA   Ledbetter D.H., Carrozzo R., Patel K., Sheer D., Lehrach H.,
RA   North M.A.;
RT   "Olfactory receptor gene cluster on human chromosome 17: possible
RT   duplication of an ancestral receptor repertoire.";
RL   Hum. Mol. Genet. 3:229-235(1994).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF095725; AAF03262.1; ALT_INIT.
DR   EMBL; AF087926; AAF37316.1; -.
DR   EMBL; U76377; AAD00250.1; -.
DR   EMBL; U78308; AAD00277.1; -.
DR   EMBL; U04688; AAA18351.1; -.
DR   GCRDb; GCR_0855; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
```

```
KW   Multigene family; Olfaction.
FT   DOMAIN        1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     29     52       1 (POTENTIAL).
FT   DOMAIN       53     60       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     61     82       2 (POTENTIAL).
FT   DOMAIN       83    103       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    104    123       3 (POTENTIAL).
FT   DOMAIN      124    142       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    144    161       4 (POTENTIAL).
FT   DOMAIN      162    199       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    200    222       5 (POTENTIAL).
FT   DOMAIN      223    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    262       6 (POTENTIAL).
FT   DOMAIN      263    275       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    276    295       7 (POTENTIAL).
FT   DOMAIN      296    315       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    100    192       BY SIMILARITY.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    165    165       A -> G (IN REF. 3).
FT   CONFLICT    183    183       D -> E (IN REF. 3).
FT   CONFLICT    187    187       L -> V (IN REF. 3).
SQ   SEQUENCE   315 AA;  34275 MW;  69B976A0FF80CCDA CRC64;
```

```
  Query Match           41.9%;  Score 346;  DB 1;  Length 315;
  Best Local Similarity 39.5%;  Pred. No. 2.6e-25;
  Matches   64;  Conservative   38;  Mismatches   60;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           :||| || |   ||  :   |:: |:   |: | |   :  | ||  | :::|||:: |
Db     127 LAICQPLTYSTRMSQTVQRMLVAASWACAFTNALTHTVAMSTLNFCGPNEVNHFYCDLPQ 186

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:::||:   :| |:::|:  ||: :  |: | :||  |:  :|: :| :|| |||||||:
Db     187 LFQLSCSSTQLNELLLFVAAAFMAVAPLVFISVSYAHVVAAVLQIRSAEGRKKAFSTCGS 246

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  | ::|||  :| |:|  |  : |:||   :| |:| |:|
Db     247 HLTVVGIFYGTGVFSYMRLGSVESSDKDKGVGVFMTVINPML 288
```

```
RESULT  15
O5V1_HUMAN
ID   O5V1_HUMAN     STANDARD;      PRT;   321 AA.
AC   Q9UGF6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 5V1 (Hs6M1-21).
GN   OR5V1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tracey A.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL096770; CAB65797.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    158       4 (POTENTIAL).
FT   DOMAIN      159    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    219       5 (POTENTIAL).
FT   DOMAIN      220    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    259       6 (POTENTIAL).
FT   DOMAIN      260    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    321       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     97    189       BY SIMILARITY.
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   321 AA;  36056 MW;  AAC426FDCF58E375 CRC64;
```

```
  Query Match           41.9%;  Score 346;  DB 1;  Length 321;
  Best Local Similarity 42.0%;  Pred. No. 2.6e-25;
  Matches   68;  Conservative   33;  Mismatches   61;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           :||||||  | |::|  |  || :  :  |||: :||  |   | ||  | |:||:|:|
Db     124 IAICNPLRYSVILSKVLCNQLAASCWAAGFLNSVVHTVLTFCLPFCGNNQINYFFCDIPP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |  :||   |:| | :   | ||     : |::||  ::  ||: :| :|| ||||||  :
Db     184 LLILSCGNTSVNELALLSTGVFIGWTPFLCIVLSYICIISTILRIQSSEGRRKAFSTCAS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||  | |:||: || |||| |  :  :|:: |: |::: |:|
Db     244 HLAIVFLFYGSAIFTYVRPISTYSLKKDRLVSVLYSVVTPML 285
```

```
Search completed: June 27, 2002, 14:11:27
Job time: 459 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:03:03 ; Search time 66.28 Seconds
(without alignments)
422.830 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 825
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	825	100.0	162	4	Q9NZP5	Q9nzp5 homo sapien
2	806	97.7	162	6	Q9N1M1	Q9n1m1 pan troglod
3	562	68.1	221	11	Q9Z231	Q9z231 rattus norv
4	527	63.9	162	6	Q9N1N4	Q9n1n4 papio hamad
5	403	48.8	268	11	Q9ERX2	Q9erx2 mus musculu
6	400	48.5	312	13	Q90807	Q90807 gallus gall
7	395	47.9	308	11	Q9EQA0	Q9eqa0 mus musculu
8	393	47.6	223	11	Q9Z1V5	Q9z1v5 mus musculu
9	393	47.6	314	11	Q9EQB8	Q9eqb8 mus musculu
10	389	47.2	310	11	Q9EQA7	Q9eqa7 mus musculu
11	387	46.9	221	11	Q9Z230	Q9z230 rattus norv
12	387	46.9	310	11	Q9EQB0	Q9eqb0 mus musculu
13	385	46.7	304	11	Q9QW36	Q9qw36 rattus sp.
14	382	46.3	157	11	Q62345	Q62345 mus musculu
15	382	46.3	309	11	Q63395	Q63395 rattus norv
16	380	46.1	158	13	O42214	O42214 necturus ma
17	380	46.1	217	4	Q96RC7	Q96rc7 homo sapien
18	380	46.1	313	6	O77758	O77758 hylobates l
19	379.5	46.0	314	11	Q9EQA5	Q9eqa5 mus musculu
20	378.5	45.9	309	11	Q9EQA6	Q9eqa6 mus musculu
21	378	45.8	162	6	Q9N1N3	Q9n1n3 papio hamad
22	378	45.8	308	11	Q9EQ98	Q9eq98 mus musculu
23	377	45.7	223	11	Q9Z1U7	Q9z1u7 mus musculu
24	377	45.7	268	11	Q9ERX3	Q9erx3 rattus norv
25	377	45.7	308	11	Q9EQ99	Q9eq99 mus musculu
26	375.5	45.5	309	11	Q9EQA4	Q9eqa4 mus musculu
27	374	45.3	311	11	Q9EQB4	Q9eqb4 mus musculu
28	374	45.3	332	13	Q90806	Q90806 gallus gall
29	373	45.2	216	4	Q96RB2	Q96rb2 homo sapien
30	372	45.1	176	6	O77545	O77545 sus scrofa
31	371.5	45.0	309	11	Q9EQA8	Q9eqa8 mus musculu
32	371	45.0	157	4	Q15617	Q15617 homo sapien
33	370	44.8	162	6	Q9N1G2	Q9n1g2 saimiri bol
34	370	44.8	216	4	Q96RC2	Q96rc2 homo sapien
35	369	44.7	311	11	Q9EQA3	Q9eqa3 mus musculu
36	369	44.7	314	6	O77756	O77756 pan troglod
37	368	44.6	318	11	Q9QY00	Q9qy00 mus musculu
38	367.5	44.5	215	4	Q96RD0	Q96rd0 homo sapien
39	367	44.5	237	11	Q9R0G3	Q9r0g3 marmota mar
40	367	44.5	312	13	Q90808	Q90808 gallus gall
41	365.5	44.3	217	4	Q96RC3	Q96rc3 homo sapien
42	364	44.1	162	6	Q9N1H0	Q9n1h0 saimiri sci
43	364	44.1	217	4	Q96RA9	Q96ra9 homo sapien
44	364	44.1	312	11	Q9EQB1	Q9eqb1 mus musculu
45	362	43.9	157	6	Q28307	Q28307 canis famil

ALIGNMENTS

RESULT 1
Q9NZP5
ID Q9NZP5 PRELIMINARY; PRT; 162 AA.
AC Q9NZP5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE OLFACTORY RECEPTOR (FRAGMENT).
GN HSA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20183981; PubMed=10706615;
RA Rouquier S., Blancher A., Giorgi D.;
RT "The olfactory receptor gene repertoire in primates and mouse:
RT Evidence for reduction of the functional fraction in primates.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR EMBL; AF179759; AAF40348.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 162 162
SQ SEQUENCE 162 AA; 18386 MW; 6716C48C3ED0E4D8 CRC64;

Query Match 100.0%; Score 825; DB 4; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.1e-77;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120

```
Db     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||||||||||||||||||||||||||||||||||||
Db    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162


RESULT   2
Q9N1M1
ID   Q9N1M1      PRELIMINARY;      PRT;   162 AA.
AC   Q9N1M1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   PTR210.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179735; AAF40330.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     162    162
SQ   SEQUENCE   162 AA;  18389 MW;  586280CC7BD5F4CE CRC64;


  Query Match             97.7%;  Score 806;  DB 6;  Length 162;
  Best Local Similarity   97.5%;  Pred. No. 1e-75;
  Matches  158;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||| |
Db      61 LFKISCNGPSINALMIFIFGAFIQIPTLMTIIISYSRVLFDILKKKSEKGRSKAFSTCSA 120

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |||||||||||||||||||||| ||||||||||||||||
Db     121 HLLSVSLYYGTLIFMYVRPASGLAEDPDKVYSLFYTIIIPLL 162


RESULT   3
Q9Z231
ID   Q9Z231      PRELIMINARY;      PRT;   221 AA.
AC   Q9Z231;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HFV-FD1;
RA   Singer M.S., Hughes T.E., Shepherd G.M., Greer C.A.;
RT   "Isolation of olfactory receptor mRNA sequences from olfactory bulb
RT   glomerular layer.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF091575; AAC64595.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     221    221
SQ   SEQUENCE   221 AA;  25000 MW;  73C05C4A5318E5DE CRC64;


  Query Match             68.1%;  Score 562;  DB 11;  Length 221;
  Best Local Similarity   65.2%;  Pred. No. 2.3e-50;
  Matches  105;  Conservative   25;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||| |:|| :|   |:|:||:||||:|:||| || |||||| |:| :|||||||
Db      61 VAICNPLLYLVVMSKRLCTVLISMSYIIGFLNPIVHVGLLFRLTFCRSNVIDHFYCEILP 120

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |: ||| |::|||:|||| : ||| | :||::|| |||| :|  |||||| ||| || :
Db     121 LYTISCTDPTLNALVIFIFASSIQISTSVTIVVSYARVLFAVLYMKSEKGRRKAFFTCSS 180

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPL 161
           || ||||:||||:|||   |   |:|||:|||||::|||
Db     181 HLFSVSLFYGTLLFMYVSHGSAPGENQDKMYSLFYTVVIPL 221


RESULT   4
Q9N1N4
ID   Q9N1N4      PRELIMINARY;      PRT;   162 AA.
AC   Q9N1N4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   PPA138.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179721; AAF40317.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     162    162
SQ   SEQUENCE   162 AA;  18112 MW;  ACD9BA1ECBD37959 CRC64;


  Query Match             63.9%;  Score 527;  DB 6;  Length 162;
  Best Local Similarity   62.3%;  Pred. No. 7e-47;
  Matches  101;  Conservative   22;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||||:|:| |  :|| :|:| |||| |:|  :| |||||  ||||:|||:|:
Db       1 VAICKPLLYPVIMTNGLCIRLLVLSFVGGFLHALIHEGILFRLTFCNSNIIHHFYCDIIP 60

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |  |||  |||| |::||    ||: |:::|:::||  ||| |||||| ||  ||||||||
Db      61 LLTISCTDPSINFLMLFILSGSIQVFTILTVLVSYAFVLFTILKKKSVKGIRKAFSTCGA 120

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           || || |||| |:|||| |||  |:||| |  :|||:||| |
```

```
Db    121 HLFSVCLYYGPLLFMYVGPASPQADDQDMVECVFYTVIIPFL 162


RESULT   5
Q9ERX2
ID   Q9ERX2      PRELIMINARY;      PRT;   268 AA.
AC   Q9ERX2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=TESTIS;
RX   MEDLINE=20469363; PubMed=11014824;
RA   Branscomb A., Seger J., White R.L.;
RT   "Evolution of odorant receptors expressed in mammalian testes.";
RL   Genetics 156:785-797(2000).
DR   EMBL; AF271049; AAG21322.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     268    268
SQ   SEQUENCE   268 AA;  29757 MW;  F7AAE1764094E78C CRC64;


  Query Match             48.8%;  Score 403;  DB 11;  Length 268;
  Best Local Similarity   42.0%;  Pred. No. 7.6e-34;
  Matches   68;  Conservative   47;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| |||| |:||::| : |: :::::||:  | | | :: |:||: :|| :::|::|
Db     100 VAICRPLLYNVIMSSRLCSLLVLVAFILGFVSALAHTSAMMNLSFCKSHIISHYFCDVLP 159

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |  :||:   :| |::|| | |  :   : : |||  :   ||  || |||||:| || :
Db     160 LLNLSCSNTHLNELLLFIIGGFNTLVPTLAVAISYVFIFCSILHIKSSKGRSKSFGTCSS 219

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::| :::|:: ||| :|:|  : :|:|| |:||| :||:|
Db     220 HLMAVGIFFGSITFMYFKPSSSNSLEQEKVSSVFYTTVIPML 261


RESULT   6
Q90807
ID   Q90807      PRELIMINARY;      PRT;   312 AA.
AC   Q90807;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   OLFACTORY RECEPTOR 3 (FRAGMENT).
GN   COR3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96184947; PubMed=8608858;
RA   Leibovici M., Lapointe F., Aletta P., Ayer-Le Lievre C.;
RT   "Avian olfactory receptors : differentiation of olfactory neurons
RT   under normal and experimental conditions.";
RL   Dev. Biol. 175:118-131(1996).
DR   EMBL; X94743; CAA64369.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER     312    312
SQ   SEQUENCE   312 AA;  35258 MW;  D629D3904184C098 CRC64;


  Query Match             48.5%;  Score 400;  DB 13;  Length 312;
  Best Local Similarity   43.8%;  Pred. No. 1.8e-33;
  Matches   71;  Conservative   47;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| ||||| :|:  :  :|:   | : ||: ||| | ||:|:|| |::::|:|:
Db     124 VAICKPLLYPSIMTKAVCWRLVKGLYFLAFLNSLVHTSGLLKLSFCSSNVVNHFFCDNSP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           ||:|| :  ::| |:: |||:   : :::||:|||  ::  ::: :|: |: ||||| :
Db     184 LFQISSSSTTLNELLVIIFGSLFVMSSIITILISYVFIILTVVRIRSKDGKYKAFSTCTS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::|||::||:||||:|     : | ||: |||||::||:|
Db     244 HLMAVSLFHGTVIFMYLRSVKLFSLDTDKIASLFYTVVIPML 285


RESULT   7
Q9EQA0
ID   Q9EQA0      PRELIMINARY;      PRT;   308 AA.
AC   Q9EQA0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K40.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282289; AAG39874.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   308 AA;  34031 MW;  6279F1C7089AA820 CRC64;


  Query Match             47.9%;  Score 395;  DB 11;  Length 308;
  Best Local Similarity   41.4%;  Pred. No. 5.9e-33;
  Matches   67;  Conservative   46;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| |||| |:||::| : |: :::::||:  | | | :: |:||: ::| :::|::|
Db     124 VAICRPLLYNVIMSSRLCSLLVLVAFILGFVSALAHTSAMMNLSFCKSHVISHYFCDVLP 183

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           |  :||:   :| |::||   |  :   : : |||  :   ||  || |||||||| || :
Db     184 LLNLSCSDIKLNELLLFIIAGFNTLVPTLAVAISYVFIFCSILHIKSSKGRSKAFGTCSS 243

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::| :::|:: ||| :| |  : :|:|| |:||| :||:|
Db     244 HLMAVGIFFGSITFMYFKPPSSNSLEQEKVSSVFYTTVIPML 285
```

```
RESULT   8
Q9Z1V5
ID   Q9Z1V5      PRELIMINARY;      PRT;   223 AA.
AC   Q9Z1V5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR B3 (FRAGMENT).
GN   OLFR44.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BL/6; TISSUE=OLFACTORY EPITHELIUM;
RX   MEDLINE=99091050; PubMed=9875846;
RA   Krautwurst D., Yau K.W., Reed R.R.;
RT   "Identification of ligands for olfactory receptors by functional
RT   expression of a receptor library.";
RL   Cell 95:917-926(1998).
DR   EMBL; AF102518; AAD13310.1; -.
DR   MGD; MGI:1333829; Olfr44.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     223    223
SQ   SEQUENCE   223 AA;  25089 MW;  83097D299076AF5E CRC64;


  Query Match             47.6%;  Score 393;  DB 11;  Length 223;
  Best Local Similarity   43.2%;  Pred. No. 6.8e-33;
  Matches   70;  Conservative   45;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
             ||||||||| | || ::   |:|  |:|| :    |   ::|: ||: ::|::::|::|
Db      61 VAICNPLLYNVTMSYQIYIFLISGVYIIGVICASAHTGFMVRIRFCKLDVINHYFCDLLP 120

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |::|:   || ::|  ||       ::||| ||  ::  ||: :| :|||||||| :
Db     121 LLKLACSNTYINEMLILFFGTLNIFVPILTIITSYIFIIASILRIRSTEGRSKAFSTCSS 180

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |:|:|::::|:| |||::|:|  : || || |:||||::|:|
Db     181 HILAVAVFFGSLAFMYLQPSSVSSMDQGKVSSVFYTIVVPML 222


RESULT   9
Q9EQB8
ID   Q9EQB8      PRELIMINARY;      PRT;   314 AA.
AC   Q9EQB8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282271; AAG39856.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   314 AA;  35440 MW;  252457E8D336CE64 CRC64;


  Query Match             47.6%;  Score 393;  DB 11;  Length 314;
  Best Local Similarity   43.2%;  Pred. No. 9.6e-33;
  Matches   70;  Conservative   45;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
             ||||||||| | || ::   |:|  |:|| :    |   ::|: ||: ::|::::|::|
Db     127 VAICNPLLYNVTMSYQIYIFLISGVYIIGVICASAHTGFMVRIRFCKLDVINHYFCDLLP 186

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |::|:   || ::|  ||       ::||| ||  ::  ||: :| :|||||||| :
Db     187 LLKLACSNTYINEMLILFFGTLNIFVPILTIITSYIFIIASILRIRSTEGRSKAFSTCSS 246

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           |:|:|::::|:| |||::|:|  : || || |:||||::|:|
Db     247 HILAVAVFFGSLAFMYLQPSSVSSMDQGKVSSVFYTIVVPML 288


RESULT  10
Q9EQA7
ID   Q9EQA7      PRELIMINARY;      PRT;   310 AA.
AC   Q9EQA7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K23.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282282; AAG39867.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   310 AA;  34770 MW;  6D627896FD82CC2F CRC64;


  Query Match             47.2%;  Score 389;  DB 11;  Length 310;
  Best Local Similarity   47.9%;  Pred. No. 2.5e-32;
  Matches   78;  Conservative   35;  Mismatches   48;  Indels    2;  Gaps    2;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
             |||||||:| | || ::   ||   |::||:  : |   : |||||  ||:::: |:||
Db     123 VAICNPLMYMVTMSPQVCLLLLFGVYLMGFVGAMAHTISMARLTFCADNIVNHYMCDILP 182

Qy      61 LFKISCNGPSINALIIFIFGAF-IQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCG 119
           | : ||    :| |::||| :| | :| ::|| |||  :|  ||   | :|||||||||
Db     183 LLEHSCTSTYVNELVVFIFVSFDIGVP-IVTIFISYALILSSILHMHSTEGRSKAFSTCS 241

Qy     120 AHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           :|:: | |::|:  |||::| | |: || || ||||||::|:|
Db     242 SHMIVVCLFFGSGAFMYLQPPSVLSLDQGKVSSLFYTIVVPML 284
```

```
RESULT  11
Q9Z230
ID   Q9Z230      PRELIMINARY;      PRT;   221 AA.
AC   Q9Z230;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HFL-VN1;
RA   Singer M.S., Hughes T.E., Shepherd G.M., Greer C.A.;
RT   "Isolation of olfactory receptor mRNA sequences from olfactory bulb
RT   glomerular layer.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF091576; AAC64596.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     221    221
SQ   SEQUENCE   221 AA;  24719 MW;  1FF1BF090F84E4C1 CRC64;


  Query Match             46.9%;  Score 387;  DB 11;  Length 221;
  Best Local Similarity   43.1%;  Pred. No. 2.8e-32;
  Matches   69;  Conservative   42;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       2 AICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQL 61
           |||||||| | || ::   |:   |: | |   ||   ::|: ||: ::|::::|::| |
Db      62 AICNPLLYNVTMSYQIYTSLILGVYIFGMLCASVHTGFMIRVKFCKLDVINHYFCDLLPL 121

Qy      62 FKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAH 121
           |::|:   || ::|  ||       |:||| ||  ::  ||  :| |||||||||| :|
Db     122 LKLACSNTYINEMLILFFGTLNICVPLLTIITSYIFIIVSILHIRSTKGRSKAFSTCSSH 181

Qy     122 LLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPL 161
           :::|::::|:| |||::|:|  : || || |:||||::|:
Db     182 IVAVAIFFGSLAFMYLQPSSVSSMDQGKVSSVFYTIVVPM 221


RESULT  12
Q9EQB0
ID   Q9EQB0      PRELIMINARY;      PRT;   310 AA.
AC   Q9EQB0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K21.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282279; AAG39864.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   310 AA;  34451 MW;  E9C6E3C269BFA6DD CRC64;


  Query Match             46.9%;  Score 387;  DB 11;  Length 310;
  Best Local Similarity   46.6%;  Pred. No. 4e-32;
  Matches   76;  Conservative   38;  Mismatches   47;  Indels    2;  Gaps    2;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||:| | || ::   ||   ||:||   : | : :::|||||   ::::: |:||
Db     123 VAICNPLMYTVTMSPQVCLLLLLGVYVMGFAGAMAHTAFMVKLTFCADKLVNHYMCDILP 182

Qy      61 LFKISCNGPSINALIIFI-FGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCG 119
           | : ||    :| |::||  |  | :|| :|| |||  :|  ||:  | :|||||||||
Db     183 LLERSCTSTYVNELVVFIVVGIDIGVPT-VTIFISYALILSSILRISSTEGRSKAFSTCS 241

Qy     120 AHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           :|:::|||::|:  |||::|:| |  :| || ||||||::|:|
Db     242 SHIIAVSLFFGSGAFMYLKPSSLLPMNQGKVSSLFYTIVVPML 284


RESULT  13
Q9QW36
ID   Q9QW36      PRELIMINARY;      PRT;   304 AA.
AC   Q9QW36;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OR14=ODORANT RECEPTOR.
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93149273; PubMed=7678922;
RA   Raming K., Krieger J., Strotmann J., Boekhoff I., Kubick S.,
RA   Baumstark C., Breer H.;
RT   "Cloning and expression of odorant receptors.";
RL   Nature 361:353-356(1993).
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   304 AA;  33855 MW;  A4AC2C66CBC9B798 CRC64;


  Query Match             46.7%;  Score 385;  DB 11;  Length 304;
  Best Local Similarity   48.5%;  Pred. No. 6.3e-32;
  Matches   79;  Conservative   32;  Mismatches   50;  Indels    2;  Gaps    2;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||||||:| | || ::   ||  :||:||   : |   |: ||||  |::::| |:||
Db     117 VAICNPLMYTVTMSPQVCLLLLLGAYVMGFSEAMAHTGNLMNLTFCADNLVNHFMCDILP 176

Qy      61 LFKISCNGPSINALIIFIFGAF-IQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCG 119
           | ::|||   || |::||  |  | :| :::| |||  :|  ||:  | :|||||||||
Db     177 LLELSCNSTFINELVVFIVVAIDIAVP-IVSIFISYALILSSILRMHSTEGRSKAFSTCS 235

Qy     120 AHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           :||: | | :|:  |||::  | |  || || ||||||::|:|
Db     236 SHLIVVCLLFGSGAFMYLKLPSILPLDQGKVSSLFYTIVVPML 278


RESULT  14
Q62345
ID   Q62345      PRELIMINARY;      PRT;   157 AA.
AC   Q62345;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   PUTATIVE OLFACTORY RECEPTOR (FRAGMENT).
GN   OLFR26 OR TPCR09.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
DR   EMBL; X89681; CAA61828.1; -.
DR   MGD; MGI:109309; Olfr26.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     157    157
SQ   SEQUENCE   157 AA;  17296 MW;  F3873CCD0A5743EB CRC64;

  Query Match             46.3%;  Score 382;  DB 11;  Length 157;
  Best Local Similarity   40.8%;  Pred. No. 6.6e-32;
  Matches   64;  Conservative   46;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       3 ICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQLF 62
           || |||| |:||::| : |: :::::||:  | | | :: |:||: :|| :::|::| |
Db       1 ICRPLLYNVIMSSRLCSLLVLVAFILGFVSALAHTSAMMNLSFCKSHIISHYFCDVLPLL 60

Qy      63 KISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGAHL 122
             :||:   :| |::|| | |  :   : : |||  :   ||  || |||||:| || :||
Db      61 NLSCSNTHLNELLLFIIGGFNTLVPTLAVAISYVFIFCSILHIKSSKGRSKSFGTCSSHL 120

Qy     123 LSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIII 159
           ::| :::|:: ||| :|:|  : :|:|| |:||| :|
Db     121 MAVGIFFGSITFMYFKPSSSNSLEQEKVSSVFYTTVI 157


RESULT  15
Q63395
ID   Q63395      PRELIMINARY;      PRT;   309 AA.
AC   Q63395;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   OLFACTORY RECEPTOR.
GN   OLP4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=OLFACTORY EPITHELIA;
RX   MEDLINE=95045546; PubMed=7957207;
RA   Gat U., Nekrasova E., Lancet D., Natochin M.;
RT   "Olfactory receptor proteins: expression, characterization and partial
RT   purification.";
RL   Eur. J. Biochem. 225:1157-1168(1994).
DR   EMBL; X80671; CAA56697.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   309 AA;  34474 MW;  51A145ACCADE0249 CRC64;

  Query Match             46.3%;  Score 382;  DB 11;  Length 309;
  Best Local Similarity   40.1%;  Pred. No. 1.3e-31;
  Matches   65;  Conservative   48;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
           |||| |||| |:||::: : |: ::: :|    :|| | :: |:||:  || :::|: |
Db     125 VAICRPLLYNVIMSSRICSLLVLVAFSLGLFSAVVHTSAMMNLSFCKSYIISHYFCDALP 184

Qy      61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
           | |::|:   :| |:||| |    :   : : ||| :   ||: :| :|||||| || :
Db     185 LLKLACSNTHLNELLIFIIGGLNTLVPTLAVAISYVFIFCSILRIRSSEGRSKAFGTCSS 244

Qy     121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
           ||::| :::|:: |||::|:|  : :|:|| |:||| :||:|
Db     245 HLMAVGIFFGSITFMYLKPSSSNSLEQEKVSSVFYTTVIPML 286


Search completed: June 27, 2002, 14:10:43
Job time: 460 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:04:48 ; Search time 30.51 Seconds
(without alignments)
589.772 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 162
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	162	100.0	162	22	AAG98518	Human olfactory re
2	162	100.0	162	22	AAG72109	Human olfactory re
3	162	100.0	162	22	AAG72421	Human OR-like poly
4	74	45.7	162	22	AAG98500	Pan troglodytes ol
5	18	11.1	314	22	AAU24633	Human olfactory re
6	18	11.1	314	22	AAG71500	Human olfactory re
7	18	11.1	314	22	AAG72474	Human OR-like poly
8	18	11.1	319	22	AAG71531	Human olfactory re
9	18	11.1	329	22	AAG71992	Human olfactory re
10	16	9.9	310	22	AAG71502	Human olfactory re
11	15	9.3	309	22	AAG71525	Human olfactory re
12	15	9.3	324	22	AAG71529	Human olfactory re
13	15	9.3	325	22	AAU24632	Human olfactory re
14	10	6.2	162	22	AAG98487	Papio hamadryas ol
15	10	6.2	312	22	AAU24584	Human olfactory re
16	10	6.2	312	22	AAG71635	Human olfactory re
17	10	6.2	316	22	AAE06769	Human G-protein co
18	10	6.2	316	22	AAG71567	Human olfactory re
19	10	6.2	316	22	AAG71946	Human olfactory re
20	10	6.2	328	22	AAG66705	Human GPCR3 polype
21	10	6.2	333	22	AAG72668	Murine OR-like pol
22	10	6.2	333	22	AAG72682	Murine OR-like pol
23	9	5.6	124	22	AAM99944	Human expressed po
24	9	5.6	124	22	AAM99971	Human expressed po
25	9	5.6	124	22	ABB10248	Human cDNA SEQ ID
26	9	5.6	124	22	AAU18107	Novel human uterin
27	9	5.6	124	22	AAU18343	Human endocrine po
28	9	5.6	124	22	AAU18665	Renal and cardiova
29	9	5.6	124	22	AAM89713	Human immune/haema
30	9	5.6	124	22	AAU21645	Novel human neopla
31	9	5.6	125	22	AAG72560	Human OR-like poly
32	9	5.6	157	22	AAG71958	Human olfactory re
33	9	5.6	157	22	AAG72041	Human olfactory re
34	9	5.6	159	22	AAG72887	Human olfactory re
35	9	5.6	161	22	AAG98574	Saimiri boliviensi
36	9	5.6	162	22	AAG98488	Papio hamadryas ol
37	9	5.6	162	22	AAG98516	Gorilla olfactory
38	9	5.6	162	22	AAG98528	Eulemar fulvus olf
39	9	5.6	162	22	AAG98540	Macaca sylvanus ol
40	9	5.6	162	22	AAG98547	Callithrix jacchus
41	9	5.6	162	22	AAG98550	Callithrix jacchus
42	9	5.6	162	22	AAG98569	Saimiri sciureus o
43	9	5.6	162	22	AAG98577	Saimiri boliviensi
44	9	5.6	162	22	AAG72423	Human OR-like poly
45	9	5.6	179	22	AAG72676	Murine OR-like pol

ALIGNMENTS

RESULT 1
AAG98518
ID AAG98518 standard; Protein; 162 AA.
XX
AC AAG98518;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human olfactory receptor 1.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance.
XX
OS Homo sapiens.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB02017.
XX
PR 22-DEC-1999; 99US-0171746.
PR 21-DEC-2000; 2000US-0747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR N-PSDB; AAH84009.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful
PT for analysis odours e.g. in food processing and perfumery -
XX

PS Claim 3; Page 266; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 162; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.9e-159;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
Db 1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
Db 61 lfkiscngpsinaliififgafiqiptlmtiiisytrvlfdilkkksekgrskafstcga 120

Qy 121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
Db 121 hllsvslyygtlifmyvrpasglaedqdkvyslfytiiipll 162

RESULT 2
AAG72109
ID AAG72109 standard; Protein; 162 AA.
XX
AC AAG72109;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1790.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 1178; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 162; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.9e-159;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
Db 1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy 61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
Db 61 lfkiscngpsinaliififgafiqiptlmtiiisytrvlfdilkkksekgrskafstcga 120

Qy 121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
Db 121 hllsvslyygtlifmyvrpasglaedqdkvyslfytiiipll 162

RESULT 3
AAG72421
ID AAG72421 standard; Protein; 162 AA.
XX
AC AAG72421;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human OR-like polypeptide query sequence, SEQ ID NO: 2102.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Example 6; Page 1411-1412; 1857pp; English.
XX
CC The present sequence is a polypeptide encoded by one of 344 newly mined
CC human genes. It was used as a query sequence in a database search of
CC olfactory receptor (OR)-like sequences. The invention relates to isolated
CC polynucleotides encoding polypeptides involved in olfactory sensation.
CC The polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary scents
CC and the identification of the odour receptors used to detect these
CC primary scents. The methods also enable determination of secondary scents

```
CC   and the identification of combinations of odour receptors that are
CC   involved in detecting such secondary scents. This enables the
CC   construction of a scent representation (also called a scent fingerprint
CC   or scent profile), which may be used to re-create and edit scents.
CC   Libraries of olfactory receptors are useful for determining the
CC   interaction pattern of a composition with the receptors, and can be
CC   used for determining differences in the olfactory faculties of different
CC   individuals.
XX
SQ   Sequence   162 AA;


  Query Match             100.0%;  Score 162;  DB 22;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-159;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 lfkiscngpsinaliififgafiqiptlmtiiisytrvlfdilkkksekgrskafstcga 120

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||||||||||||||||||||||||||||||||||||
Db    121 hllsvslyygtlifmyvrpasglaedqdkvyslfytiiipll 162


RESULT   4
AAG98500
ID   AAG98500 standard; Protein; 162 AA.
XX
AC   AAG98500;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Pan troglodytes olfactory receptor 9.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance.
XX
OS   Pan troglodytes.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   N-PSDB; AAH83985.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 3; Page 226; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence   162 AA;


  Query Match             45.7%;  Score 74;  DB 22;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-68;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 vaicnpllypvmmsnklsaqllsisyvigflhplvhvslllrltfcrfniihyfyceilq 60

Qy     61 LFKISCNGPSINAL 74
          ||||||||||||||
Db     61 lfkiscngpsinal 74


RESULT   5
AAU24633
ID   AAU24633 standard; Protein; 314 AA.
XX
AC   AAU24633;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human olfactory receptor AOLFR127.
XX
KW   Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW   food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS   Homo sapiens.
XX
PN   WO200168805-A2.
XX
PD   20-SEP-2001.
XX
PF   13-MAR-2001; 2001WO-US07771.
XX
PR   13-MAR-2000; 2000US-0188914.
PR   24-MAR-2000; 2000US-0192033.
PR   12-APR-2000; 2000US-0198474.
PR   24-APR-2000; 2000US-0199335.
PR   26-MAY-2000; 2000US-0207702.
PR   23-JUN-2000; 2000US-0213849.
PR   16-AUG-2000; 2000US-0226534.
PR   07-SEP-2000; 2000US-0230732.
PR   07-FEB-2001; 2001US-0266862.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zozulya S;
XX
DR   WPI; 2001-570867/64.
DR   N-PSDB; AAS42326.
XX
PT   Nucleic acids encoding human olfactory G protein-coupled receptors,
PT   useful for screening for compounds involved in olfactory sensation,
PT   where the compounds can be used in the food, pharmaceutical and
PT   cosmetic industries to customise odours -
XX
PS   Claim 60; Page 136; 319pp; English.
XX
CC   The invention relates to nucleic acids encoding human olfactory
CC   receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC   specifically recognise molecules, odourants, that elicit specific
CC   olfactory sensation. The human olfactory receptors and polynucleotides
CC   encoding them are useful for screening a library of chemical compounds
CC   for compounds that are involved in olfactory sensation. Modulators of
CC   their activity are useful for pharmacological and genetic modulation of
CC   olfactory signalling pathways. Therefore, they can be used in the food,
CC   pharmaceutical and cosmetic industries to customise odours and
CC   fragrances. The present sequence is a human olfactory receptor of the
CC   invention.
XX
```

```
SQ   Sequence   314 AA;

  Query Match            11.1%;  Score 18;  DB 22;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-10;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    113 KAFSTCGAHLLSVSLYYG 130
          ||||||||||||||||||
Db    241 kafstcgahllsvslyyg 258


RESULT   6
AAG71500
ID   AAG71500 standard; Protein; 314 AA.
XX
AC   AAG71500;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polypeptide, SEQ ID NO: 1181.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 11; Page 697-698; 1857pp; English.
XX
CC   The present sequence is an olfactory receptor which is encoded by
CC   one of a number of novel polynucleotides. The polynucleotides can be
CC   used in screening for olfactory agonists and antagonists. The methods
CC   allow for the determination of primary scents and the identification
CC   of the odour receptors used to detect these primary scents. The methods
CC   also enable determination of secondary scents and the identification of
CC   combinations of odour receptors that are involved in detecting such
CC   secondary scents. This enables the construction of a scent representation
CC   (also called a scent fingerprint or scent profile), which may be used to
CC   re-create and edit scents. Libraries of olfactory receptors are useful
CC   for determining the interaction pattern of a composition with the
CC   receptors, and can be used for determining differences in the olfactory
CC   faculties of different individuals.
XX
SQ   Sequence   314 AA;

  Query Match            11.1%;  Score 18;  DB 22;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-10;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    113 KAFSTCGAHLLSVSLYYG 130
          ||||||||||||||||||
Db    241 kafstcgahllsvslyyg 258


RESULT   7
AAG72474
ID   AAG72474 standard; Protein; 314 AA.
XX
AC   AAG72474;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human OR-like polypeptide query sequence, SEQ ID NO: 2155.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Example 6; Page 1448-1449; 1857pp; English.
XX
CC   The present sequence is a polypeptide encoded by one of 344 newly mined
CC   human genes. It was used as a query sequence in a database search of
CC   olfactory receptor (OR)-like sequences. The invention relates to isolated
CC   polynucleotides encoding polypeptides involved in olfactory sensation.
CC   The polynucleotides can be used in screening for olfactory agonists and
CC   antagonists. The methods allow for the determination of primary scents
CC   and the identification of the odour receptors used to detect these
CC   primary scents. The methods also enable determination of secondary scents
CC   and the identification of combinations of odour receptors that are
CC   involved in detecting such secondary scents. This enables the
CC   construction of a scent representation (also called a scent fingerprint
CC   or scent profile), which may be used to re-create and edit scents.
CC   Libraries of olfactory receptors are useful for determining the
CC   interaction pattern of a composition with the receptors, and can be
CC   used for determining differences in the olfactory faculties of different
CC   individuals.
XX
SQ   Sequence   314 AA;

  Query Match            11.1%;  Score 18;  DB 22;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-10;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    113 KAFSTCGAHLLSVSLYYG 130
          ||||||||||||||||||
Db    241 kafstcgahllsvslyyg 258


RESULT   8
AAG71531
ID   AAG71531 standard; Protein; 319 AA.
XX
AC   AAG71531;
```

XX
DT 31-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1212.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 722; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 319 AA;

Query Match 11.1%; Score 18; DB 22; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.5e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 KAFSTCGAHLLSVSLYYG 130
||||||||||||||||||
Db 236 kafstcgahllsvslyyg 253

RESULT 9
AAG71992
ID AAG71992 standard; Protein; 329 AA.
XX
AC AAG71992;
XX
DT 31-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1673.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 1092-1093; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 329 AA;

Query Match 11.1%; Score 18; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 NIIHYFYCEILQLFKISC 66
||||||||||||||||||
Db 172 niihyfyceilqlfkisc 189

RESULT 10
AAG71502
ID AAG71502 standard; Protein; 310 AA.
XX
AC AAG71502;
XX
DT 31-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1183.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX

PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 698-699; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 310 AA;

Query Match 9.9%; Score 16; DB 22; Length 310;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 KAFSTCGAHLLSVSLY 128
||||||||||||||||
Db 231 kafstcgahllsvsly 246

RESULT 11
AAG71525
ID AAG71525 standard; Protein; 309 AA.
XX
AC AAG71525;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1206.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 717 and 684; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 309 AA;

Query Match 9.3%; Score 15; DB 22; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 STCGAHLLSVSLYYG 130
|||||||||||||||
Db 239 stcgahllsvslyyg 253

RESULT 12
AAG71529
ID AAG71529 standard; Protein; 324 AA.
XX
AC AAG71529;
XX
DT 31-JUL-2001 (first entry)
XX
DE Human olfactory receptor polypeptide, SEQ ID NO: 1210.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 11; Page 720-721; 1857pp; English.
XX
CC The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation

CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
SQ Sequence 324 AA;

Query Match 9.3%; Score 15; DB 22; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 STCGAHLLSVSLYYG 130
||||||||||||||||
Db 239 stcgahllsvslyyg 253

RESULT 13
AAU24632
ID AAU24632 standard; Protein; 325 AA.
XX
AC AAU24632;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human olfactory receptor AOLFR126.
XX
KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS Homo sapiens.
XX
PN WO200168805-A2.
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07771.
XX
PR 13-MAR-2000; 2000US-0188914.
PR 24-MAR-2000; 2000US-0192033.
PR 12-APR-2000; 2000US-0198474.
PR 24-APR-2000; 2000US-0199335.
PR 26-MAY-2000; 2000US-0207702.
PR 23-JUN-2000; 2000US-0213849.
PR 16-AUG-2000; 2000US-0226534.
PR 07-SEP-2000; 2000US-0230732.
PR 07-FEB-2001; 2001US-0266862.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR N-PSDB; AAS42325.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors,
PT useful for screening for compounds involved in olfactory sensation,
PT where the compounds can be used in the food, pharmaceutical and
PT cosmetic industries to customise odours -
XX
PS Claim 60; Page 136; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence is a human olfactory receptor of the
CC invention.
XX
SQ Sequence 325 AA;

Query Match 9.3%; Score 15; DB 22; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 STCGAHLLSVSLYYG 130
||||||||||||||||
Db 255 stcgahllsvslyyg 269

RESULT 14
AAG98487
ID AAG98487 standard; Protein; 162 AA.
XX
AC AAG98487;
XX
DT 25-SEP-2001 (first entry)
XX
DE Papio hamadryas olfactory receptor 12.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance.
XX
OS Papio hamadryas.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB02017.
XX
PR 22-DEC-1999; 99US-0171746.
PR 21-DEC-2000; 2000US-0747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR N-PSDB; AAH83971.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful
PT for analysis odours e.g. in food processing and perfumery -
XX
PS Claim 3; Page 199; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 162 AA;

Query Match 6.2%; Score 10; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 KAFSTCGAHL 122
||||||||||
Db 113 kafstcgahl 122

RESULT 15
AAU24584
ID AAU24584 standard; Protein; 312 AA.

XX
AC AAU24584;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human olfactory receptor AOLFR74.
XX
KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS Homo sapiens.
XX
PN WO200168805-A2.
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07771.
XX
PR 13-MAR-2000; 2000US-0188914.
PR 24-MAR-2000; 2000US-0192033.
PR 12-APR-2000; 2000US-0198474.
PR 24-APR-2000; 2000US-0199335.
PR 26-MAY-2000; 2000US-0207702.
PR 23-JUN-2000; 2000US-0213849.
PR 16-AUG-2000; 2000US-0226534.
PR 07-SEP-2000; 2000US-0230732.
PR 07-FEB-2001; 2001US-0266862.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR N-PSDB; AAS42277.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors,
PT useful for screening for compounds involved in olfactory sensation,
PT where the compounds can be used in the food, pharmaceutical and
PT cosmetic industries to customise odours -
XX
PS Claim 60; Page 114; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence is a human olfactory receptor of the
CC invention.
XX
SQ Sequence 312 AA;

Query Match 6.2%; Score 10; DB 22; Length 312;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 KAFSTCGAHL 122
||||||||||
Db 235 kafstcgahl 244

Search completed: June 27, 2002, 14:12:19
Job time: 451 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:08:18 ; Search time 15.87 Seconds
(without alignments)
249.335 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 162
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7	4.3	247	1	US-08-465-980-3	Sequence 3, Appli
2	7	4.3	247	2	US-09-053-303-3	Sequence 3, Appli
3	7	4.3	247	5	PCT-US95-07093-3	Sequence 3, Appli
4	7	4.3	284	1	US-08-118-270-67	Sequence 67, Appl
5	7	4.3	284	5	PCT-US93-08528-67	Sequence 67, Appl
6	7	4.3	314	3	US-08-988-876-7	Sequence 7, Appli
7	7	4.3	504	2	US-08-466-589-4	Sequence 4, Appli
8	7	4.3	504	2	US-08-700-636-4	Sequence 4, Appli
9	7	4.3	504	3	US-08-467-574-4	Sequence 4, Appli
10	7	4.3	504	4	US-09-217-345-4	Sequence 4, Appli
11	7	4.3	541	4	US-09-158-767-19	Sequence 19, Appl
12	7	4.3	541	4	US-09-158-767-20	Sequence 20, Appl
13	7	4.3	584	1	US-08-448-196A-7	Sequence 7, Appli
14	7	4.3	606	2	US-08-883-534-3	Sequence 3, Appli
15	7	4.3	606	3	US-09-204-764-3	Sequence 3, Appli
16	7	4.3	1302	1	US-08-232-537-2	Sequence 2, Appli
17	7	4.3	4551	3	US-09-320-878-1	Sequence 1, Appli
18	7	4.3	4613	4	US-09-105-537-31	Sequence 31, Appl
19	7	4.3	5215	4	US-09-105-537-2	Sequence 2, Appli
20	7	4.3	11877	4	US-09-105-537-6	Sequence 6, Appli
21	6	3.7	13	4	US-09-461-697-15	Sequence 15, Appl
22	6	3.7	16	1	US-08-196-630A-9	Sequence 9, Appli
23	6	3.7	22	1	US-08-196-630A-8	Sequence 8, Appli
24	6	3.7	37	2	US-08-942-423-5	Sequence 5, Appli
25	6	3.7	43	2	US-08-389-360-2	Sequence 2, Appli
26	6	3.7	43	3	US-09-038-328-2	Sequence 2, Appli
27	6	3.7	67	2	US-08-639-857-31	Sequence 31, Appl
28	6	3.7	70	1	US-08-466-033-185	Sequence 185, App
29	6	3.7	70	2	US-08-444-733-185	Sequence 185, App
30	6	3.7	70	2	US-08-464-134-185	Sequence 185, App
31	6	3.7	70	2	US-08-461-361-185	Sequence 185, App
32	6	3.7	70	2	US-08-485-910-185	Sequence 185, App
33	6	3.7	81	2	US-08-332-562A-86	Sequence 86, Appl
34	6	3.7	96	1	US-08-488-113B-158	Sequence 158, App
35	6	3.7	96	1	US-08-477-484B-158	Sequence 158, App
36	6	3.7	96	1	US-08-107-669D-22	Sequence 22, Appl
37	6	3.7	96	1	US-08-472-788A-22	Sequence 22, Appl
38	6	3.7	96	2	US-08-477-531B-22	Sequence 22, Appl
39	6	3.7	96	2	US-08-646-360-158	Sequence 158, App
40	6	3.7	96	2	US-08-082-842A-22	Sequence 22, Appl
41	6	3.7	96	4	US-08-839-765-158	Sequence 158, App
42	6	3.7	96	4	US-09-136-389-158	Sequence 158, App
43	6	3.7	131	1	US-08-466-033-179	Sequence 179, App
44	6	3.7	131	2	US-08-444-733-179	Sequence 179, App
45	6	3.7	131	2	US-08-464-134-179	Sequence 179, App

ALIGNMENTS

```
RESULT 1
US-08-465-980-3
; Sequence 3, Application US/08465980
; Patent No. 5756309
;  GENERAL INFORMATION:
;    APPLICANT:  Soppet, Daniel R.
;    APPLICANT:  Li, Yi
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR HPRAJ70
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,980
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-446
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-465-980-3


  Query Match              4.3%;  Score 7;  DB 1;  Length 247;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
```

```
          |||||||
Db    225 KAFSTCG 231


RESULT   2
US-09-053-303-3
; Sequence 3, Application US/09053303
; Patent No. 5948890
;  GENERAL INFORMATION:
;    APPLICANT:  Soppet, Daniel R.
;    APPLICANT:  Li, Yi
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR HPRAJ70
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/053,303
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/465;980
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-446
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-053-303-3


  Query Match             4.3%;  Score 7;  DB 2;  Length 247;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     225 KAFSTCG 231


RESULT   3
PCT-US95-07093-3
; Sequence 3, Application PC/TUS9507093
;  GENERAL INFORMATION:
;    APPLICANT:  Soppet, Daniel R.
;    APPLICANT:  Li, Yi
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR HPRAJ70
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STUART & OLSTEIN
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07093
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/465,980
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-446
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-07093-3


  Query Match             4.3%;  Score 7;  DB 5;  Length 247;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     225 KAFSTCG 231


RESULT   4
US-08-118-270-67
; Sequence 67, Application US/08118270
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Townsend, Kevin G.
;       REGISTRATION NUMBER:  34,033
;       REFERENCE/DOCKET NUMBER:  MURPHY=2A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  67:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  284 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-118-270-67


  Query Match             4.3%;  Score 7;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     208 KAFSTCG 214


RESULT   5
PCT-US93-08528-67
; Sequence 67, Application PC/TUS9308528
;  GENERAL INFORMATION:
;     APPLICANT:  New York University
;     TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;     TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;     NUMBER OF SEQUENCES:  348
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BROWDY AND NEIMARK
;       STREET:  419 Seventh Street, N.W., Suite 300
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  20004
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/08528
;       FILING DATE:  09-SEP-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/943,236
;       FILING DATE:  10-SEP-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Townsend, Kevin G.
;       REGISTRATION NUMBER:  34,033
;       REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  67:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  284 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
PCT-US93-08528-67


  Query Match             4.3%;  Score 7;  DB 5;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     208 KAFSTCG 214


RESULT   6
US-08-988-876-7
; Sequence 7, Application US/08988876
; Patent No. 6063596
;  GENERAL INFORMATION:
;     APPLICANT:  Lal, Preeti
;     APPLICANT:  Bandman, Olga
;     APPLICANT:  Hillman, Jennifer L.
;     APPLICANT:  Yue, Henry
;     TITLE OF INVENTION:  G PROTEIN COUPLED RECEPTORS ASSOCIATED
;     TITLE OF INVENTION:  WITH IMMUNE RESPONSE
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;       STREET:  3174 Porter Drive
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/988,876
;       FILING DATE:  Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Billings, Lucy J.
;       REGISTRATION NUMBER:  36,749
;       REFERENCE/DOCKET NUMBER:  PF-0441 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-855-0555
;       TELEFAX:  650-845-4166
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  314 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  GenBank
;       CLONE:  32086
US-08-988-876-7


  Query Match             4.3%;  Score 7;  DB 3;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     236 KAFSTCG 242


RESULT   7
US-08-466-589-4
; Sequence 4, Application US/08466589
; Patent No. 5837489
;  GENERAL INFORMATION:
;     APPLICANT:  Elliot, Kathryn J.
;     APPLICANT:  Ellis, Steven B.
```

```
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClaim
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,589
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  March 8, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-9950
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-466-589-4


  Query Match             4.3%;  Score 7;  DB 2;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     152 SLFYTII 158
           |||||||
Db     240 SLFYTII 246


RESULT   8
US-08-700-636-4
; Sequence 4, Application US/08700636
; Patent No. 5910582
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,636
;      FILING DATE:  16-JUL-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  08-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9368
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-4737
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-700-636-4


  Query Match             4.3%;  Score 7;  DB 2;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     152 SLFYTII 158
           |||||||
Db     240 SLFYTII 246


RESULT   9
US-08-467-574-4
; Sequence 4, Application US/08467574
; Patent No. 6022704
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClaim
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,574
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  March 8, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-9949
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-467-574-4


  Query Match             4.3%;  Score 7;  DB 3;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     152 SLFYTII 158
           |||||||
Db     240 SLFYTII 246


RESULT  10
US-09-217-345-4
; Sequence 4, Application US/09217345
; Patent No. 6303753
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Heller Ehrman White & McAuliffe
;      STREET:  4250 Executive Square, 7th Floor
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,345
;      FILING DATE:  21-DEC-98
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/467,574
;      FILING DATE:  05-JUN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/466,589,
;      FILING DATE:  05-JUN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  08-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  24735-9949B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-450-8400
;      TELEFAX:  619-587-5360
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-09-217-345-4


  Query Match             4.3%;  Score 7;  DB 4;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     152 SLFYTII 158
           |||||||
Db     240 SLFYTII 246


RESULT  11
US-09-158-767-19
; Sequence 19, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
;  APPLICANT: Batard, Yannick
;  APPLICANT: Durst, Francis
;  APPLICANT: Schalk, Michel
;  APPLICANT: Werck-Reichhart, Daniele
;  TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
;  TITLE OF INVENTION: EXPRESSION IN YEAST AND OBTAINED TRANSFORMED YEAST
;  FILE REFERENCE: A32000
;  CURRENT APPLICATION NUMBER: US/09/158,767A
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: FR 97-12094
;  EARLIER FILING DATE: 1997-09-24
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 541
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Altered sequences
US-09-158-767-19


  Query Match             4.3%;  Score 7;  DB 4;  Length 541;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      39 LLLRLTF 45
           |||||||
Db     178 LLLRLTF 184


RESULT  12
US-09-158-767-20
; Sequence 20, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
;  APPLICANT: Batard, Yannick
;  APPLICANT: Durst, Francis
;  APPLICANT: Schalk, Michel
;  APPLICANT: Werck-Reichhart, Daniele
;  TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
;  TITLE OF INVENTION: EXPRESSION IN YEAST AND OBTAINED TRANSFORMED YEAST
;  FILE REFERENCE: A32000
;  CURRENT APPLICATION NUMBER: US/09/158,767A
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: FR 97-12094
;  EARLIER FILING DATE: 1997-09-24
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
;   LENGTH: 541
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Altered sequences
US-09-158-767-20


  Query Match             4.3%;  Score 7;  DB 4;  Length 541;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      39 LLLRLTF 45
           |||||||
Db     178 LLLRLTF 184


RESULT  13
US-08-448-196A-7
; Sequence 7, Application US/08448196A
; Patent No. 5780594
;  GENERAL INFORMATION:
;    APPLICANT:  CARTER, DANIEL C.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
;    TITLE OF INVENTION:  CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
;    TITLE OF INVENTION:  RELATED PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NASA
;      STREET:  MARSHALL SPACE FLIGHT CENTER
;      CITY:  HUNTSVILLE
;      STATE:  ALABAMA
;      COUNTRY:  USA
;      ZIP:  35812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,196A
;      FILING DATE:  23-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROAD JR., ROBERT L.
;      REGISTRATION NUMBER:  18,757
;      REFERENCE/DOCKET NUMBER:  XX/MFS-28402-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  205-544-0021
;      TELEFAX:  205-544-0258
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  584 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-448-196A-7


  Query Match             4.3%;  Score 7;  DB 1;  Length 584;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      37 VSLLLRL 43
           |||||||
Db     343 VSLLLRL 349


RESULT  14
US-08-883-534-3
; Sequence 3, Application US/08883534
; Patent No. 5846777
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  TWO NEW WD-40 PROTEINS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,534
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0332 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  606 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  NEUTGMT01
;      CLONE:  1221143
US-08-883-534-3


  Query Match             4.3%;  Score 7;  DB 2;  Length 606;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     121 HLLSVSL 127
           |||||||
Db     291 HLLSVSL 297


RESULT  15
US-09-204-764-3
; Sequence 3, Application US/09204764
; Patent No. 6025464
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  TWO NEW WD-40 PROTEINS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/204,764
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/883,534
;      FILING DATE:
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0332 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  606 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  NEUTGMT01
;      CLONE:  1221143
US-09-204-764-3


  Query Match             4.3%;  Score 7;  DB 3;  Length 606;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     121 HLLSVSL 127
           |||||||
Db     291 HLLSVSL 297


Search completed: June 27, 2002, 14:12:56
Job time: 278 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:09:18 ; Search time 18.79 Seconds
(without alignments)
828.444 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 162
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9	5.6	157	2	S58038	probable olfactory
2	9	5.6	157	2	S57998	probable olfactory
3	9	5.6	157	2	S58017	probable olfactory
4	9	5.6	157	2	S58014	probable olfactory
5	9	5.6	157	2	S58007	probable olfactory
6	9	5.6	222	2	C40745	odorant receptor (
7	9	5.6	304	2	S29709	olfactory receptor
8	8	4.9	136	2	PN0529	G protein-coupled
9	8	4.9	333	1	C69812	ferrichrome ABC tr
10	8	4.9	786	2	A69927	ribonucleoside-dip
11	8	4.9	1084	2	T12925	ribunucleotide red
12	7	4.3	81	1	W5WL35	E5 protein - human
13	7	4.3	132	2	D90362	hypothetical prote
14	7	4.3	135	2	PN0530	G protein-coupled
15	7	4.3	135	2	PN0527	G protein-coupled
16	7	4.3	147	2	T17958	hypothetical prote
17	7	4.3	157	2	S58031	probable olfactory
18	7	4.3	157	2	S58004	probable olfactory
19	7	4.3	157	2	S57995	probable olfactory
20	7	4.3	157	2	S58001	probable olfactory
21	7	4.3	157	2	S58013	probable olfactory
22	7	4.3	157	2	S58067	probable olfactory
23	7	4.3	157	2	S58030	probable olfactory
24	7	4.3	157	2	S58018	probable olfactory
25	7	4.3	157	2	S57996	probable olfactory
26	7	4.3	185	2	S28996	G protein-coupled
27	7	4.3	204	2	AH2119	hypothetical prote
28	7	4.3	216	2	I38470	olfactory receptor
29	7	4.3	216	2	I38474	olfactory receptor
30	7	4.3	216	2	I38482	olfactory receptor
31	7	4.3	216	2	I38483	olfactory receptor
32	7	4.3	216	2	I38480	olfactory receptor
33	7	4.3	216	2	I38479	olfactory receptor
34	7	4.3	216	2	I38476	olfactory receptor
35	7	4.3	216	2	I38477	olfactory receptor
36	7	4.3	216	2	I38484	olfactory receptor
37	7	4.3	217	2	T28343	hypothetical prote
38	7	4.3	222	2	B40745	odorant receptor (
39	7	4.3	225	2	I38478	olfactory receptor
40	7	4.3	232	2	C72609	hypothetical prote
41	7	4.3	234	2	S28999	G protein-coupled
42	7	4.3	234	2	S28998	G protein-coupled
43	7	4.3	264	2	PC4369	olfactory receptor
44	7	4.3	265	2	I46986	albumin - dog (fra
45	7	4.3	275	2	S47325	myoD protein - zeb

ALIGNMENTS

RESULT 1
S58038
probable olfactory receptor tpcr21 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58038
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific repe
A;Reference number: S57995
A;Accession: S58038
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89704; NID:g902724; PIDN:CAA61851.1; PID:g902725
C;Superfamily: olfactory receptor OR14

Query Match 5.6%; Score 9; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    110 GRSKAFSTC 118
          |||||||||
Db    108 GRSKAFSTC 116
```

RESULT 2
S57998
probable olfactory receptor tpcr79 - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S57998
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific repe
A;Reference number: S57995
A;Accession: S57998
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89665; NID:g902198; PIDN:CAA61812.1; PID:g902199
C;Superfamily: olfactory receptor OR14

Query Match 5.6%; Score 9; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    110 GRSKAFSTC 118
          |||||||||
Db    108 GRSKAFSTC 116
```

```
RESULT   3
S58017
probable olfactory receptor tpcr25 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58017
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58017
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89672; NID:g902327; PIDN:CAA61819.1; PID:g902328
C;Superfamily: olfactory receptor OR14

  Query Match             5.6%;  Score 9;  DB 2;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GRSKAFSTC 118
           |||||||||
Db     108 GRSKAFSTC 116


RESULT   4
S58014
probable olfactory receptor tpcr120 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58014
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58014
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89669; NID:g902321; PIDN:CAA61816.1; PID:g902322
C;Superfamily: olfactory receptor OR14

  Query Match             5.6%;  Score 9;  DB 2;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GRSKAFSTC 118
           |||||||||
Db     108 GRSKAFSTC 116


RESULT   5
S58007
probable olfactory receptor tpcr85 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C;Accession: S58007
R;Vanderhaeghen, P.; Schurmans, S.; Vassart, G.; Parmentier, M.
submitted to the EMBL Data Library, July 1995
A;Description: Male germ cells from several mammalian species express a specific reperto
A;Reference number: S57995
A;Accession: S58007
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-157 <VAN>
A;Cross-references: EMBL:X89675; NID:g902333; PIDN:CAA61822.1; PID:g902334
C;Superfamily: olfactory receptor OR14

  Query Match             5.6%;  Score 9;  DB 2;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GRSKAFSTC 118
           |||||||||
Db     108 GRSKAFSTC 116


RESULT   6
C40745
odorant receptor (clone K7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: C40745
R;Ressler, K.J.; Sullivan, S.L.; Buck, L.B.
Cell 73, 597-609, 1993
A;Title: A zonal organization of odorant receptor gene expression in the olfactory ep
A;Reference number: A40745; MUID:93258822
A;Accession: C40745
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-222 <RES>
A;Cross-references: GB:L14569; NID:g293759; PIDN:AAA39853.1; PID:g293760
A;Experimental source: olfactory epithelium
A;Note: sequence extracted from NCBI backbone (NCBIP:131750)
C;Superfamily: olfactory receptor OR14
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             5.6%;  Score 9;  DB 2;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VAICNPLLY 9
           |||||||||
Db      64 VAICNPLLY 72


RESULT   7
S29709
olfactory receptor OR14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 04-Sep-1998
C;Accession: S29709
R;Raming, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Br
Nature 361, 353-356, 1993
A;Title: Cloning and expression of odorant receptors.
A;Reference number: S29707; MUID:93149273
A;Accession: S29709
A;Molecule type: mRNA
A;Residues: 1-304 <RAM>
C;Superfamily: olfactory receptor OR14

  Query Match             5.6%;  Score 9;  DB 2;  Length 304;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GRSKAFSTC 118
           |||||||||
Db     226 GRSKAFSTC 234


RESULT   8
PN0529
G protein-coupled receptor type A (clone TAS 7) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 03-Nov-2000
C;Accession: PN0529
R;Matsuoka, I.; Mori, T.; Aoki, J.; Sato, T.; Kurihara, K.
Biochem. Biophys. Res. Commun. 194, 504-511, 1993
```

A;Title: Identification of novel members of G-protein coupled receptor superfamily expre
A;Reference number: JN0621; MUID:93326166
A;Accession: PN0529
A;Molecule type: mRNA
A;Residues: 1-136 <MAT>
A;Experimental source: tongue taste papillae
C;Comment: This protein is involved in modulating taste sensitivity or regeneration of t
C;Superfamily: olfactory receptor OR14
C;Keywords: receptor; transmembrane protein
F;1-8/Domain: transmembrane #status predicted <TM1>
F;28-46/Domain: transmembrane #status predicted <TM2>
F;90-106/Domain: transmembrane #status predicted <TM3>
F;125-135/Domain: transmembrane #status predicted <TM4>

Query Match 4.9%; Score 8; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AICNPLLY 9
||||||||
Db 13 AICNPLLY 20

RESULT 9
C69812
ferrichrome ABC transporter (permease) homolog yfmD - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: C69812
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033
A;Accession: C69812
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-333 <KUN>
A;Cross-references: GB:Z99108; GB:AL009126; NID:g2633055; PIDN:CAB12580.1; PID:g2633075
A;Experimental source: strain 168
C;Genetics:
A;Gene: yfmD
C;Superfamily: ferrichrome ABC transporter

Query Match 4.9%; Score 8; DB 1; Length 333;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 ALIIFIFG 80
||||||||
Db 19 ALIIFIFG 26

RESULT 10
A69927
ribonucleoside-diphosphate reductase (alph) homolog yosO - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: A69927
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033
A;Accession: A69927
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-786 <KUN>
A;Cross-references: GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13897.1; PID:e11854
A;Experimental source: strain 168
C;Genetics:
A;Gene: yosO

Query Match 4.9%; Score 8; DB 2; Length 786;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 VLFDILKK 105
||||||||
Db 279 VLFDILKK 286

RESULT 11
T12925
ribunucleotide reductase large chain - Bacillus subtilis phage SPBc2
C;Species: Bacillus subtilis phage SPBc2
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 02-Jun-2000
C;Accession: T12925
R;Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
A;Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 pro
A;Reference number: Z17583
A;Accession: T12925
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1084 <LAZ>
A;Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025639; PIDN:AAC13134.1
C;Genetics:
A;Gene: bnrdE

Query Match 4.9%; Score 8; DB 2; Length 1084;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 VLFDILKK 105
||||||||
Db 577 VLFDILKK 584

RESULT 12
W5WL35
E5 protein - human papillomavirus type 35
C;Species: human papillomavirus type 35
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C;Accession: D40824
R;Marich, J.E.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.
Virology 186, 770-776, 1992
A;Title: The phylogenetic relationship and complete nucleotide sequence of human papi
A;Reference number: A40824; MUID:92124753

A;Accession: D40824
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-81 <MAR>
A;Cross-references: GB:M74117; NID:g333050; PIDN:AAA46970.1; PID:g333056
C;Superfamily: papillomavirus E5 protein
C;Keywords: early protein

Query Match 4.3%; Score 7; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 122 LLSVSLY 128
|||||||
Db 32 LLSVSLY 38

RESULT 13
D90362
hypothetical protein SSO1964 [imported] - Sulfolobus solfataricus transposon ISC1913
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: D90362
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: D90362
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-132 <KUR>
A;Cross-references: GB:AE006641; NID:g13815239; PIDN:AAK42155.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO1964

Query Match 4.3%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QLLSISY 26
|||||||
Db 12 QLLSISY 18

RESULT 14
PN0530
G protein-coupled receptor type A (clone TAS 38) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 19-May-2000
C;Accession: PN0530
R;Matsuoka, I.; Mori, T.; Aoki, J.; Sato, T.; Kurihara, K.
Biochem. Biophys. Res. Commun. 194, 504-511, 1993
A;Title: Identification of novel members of G-protein coupled receptor superfamily expre
A;Reference number: JN0621; MUID:93326166
A;Accession: PN0530
A;Molecule type: mRNA
A;Residues: 1-135 <MAT>
A;Experimental source: tongue taste papillae
C;Comment: This protein is involved in modulating taste sensitivity or regeneration of t
C;Superfamily: olfactory receptor OR14
C;Keywords: receptor; transmembrane protein
F;1-8/Domain: transmembrane #status predicted <TM1>
F;28-46/Domain: transmembrane #status predicted <TM2>
F;91-106/Domain: transmembrane #status predicted <TM3>
F;125-135/Domain: transmembrane #status predicted <TM4>

Query Match 4.3%; Score 7; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAICNPL 7
|||||||
Db 12 VAICNPL 18

RESULT 15
PN0527
G protein-coupled receptor type A (clone TAS 4) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 03-Nov-2000
C;Accession: PN0527
R;Matsuoka, I.; Mori, T.; Aoki, J.; Sato, T.; Kurihara, K.
Biochem. Biophys. Res. Commun. 194, 504-511, 1993
A;Title: Identification of novel members of G-protein coupled receptor superfamily ex
A;Reference number: JN0621; MUID:93326166
A;Accession: PN0527
A;Molecule type: DNA
A;Residues: 1-135 <MAT>
A;Experimental source: tongue taste papillae
A;Note: cross-reference
C;Comment: This protein is involved in modulating taste sensitivity or regeneration o
C;Superfamily: olfactory receptor OR14
C;Keywords: receptor; transmembrane protein
F;1-8/Domain: transmembrane #status predicted <TM1>
F;28-46/Domain: transmembrane #status predicted <TM2>
F;90-106/Domain: transmembrane #status predicted <TM3>
F;125-135/Domain: transmembrane #status predicted <TM4>

Query Match 4.3%; Score 7; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 KAFSTCG 119
|||||||
Db 124 KAFSTCG 130

Search completed: June 27, 2002, 14:13:35
Job time: 257 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:12:58 ; Search time 10.69 Seconds
(without alignments)
586.770 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 162
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9	5.6	159	1	OLF8_CHICK	Q98913 gallus gall
2	9	5.6	222	1	OL7B_MOUSE	P34983 mus musculu
3	9	5.6	311	1	O8B8_HUMAN	Q15620 homo sapien
4	9	5.6	311	1	OLF1_CANFA	Q95154 canis famil
5	9	5.6	314	1	O5I1_HUMAN	Q13606 homo sapien
6	7	4.3	81	1	VE5_HPV35	P27226 human papil
7	7	4.3	111	1	OL7H_MOUSE	Q60893 mus musculu
8	7	4.3	112	1	OL7C_MOUSE	Q60882 mus musculu
9	7	4.3	112	1	OL7D_MOUSE	Q60884 mus musculu
10	7	4.3	185	1	GU01_RAT	P35894 rattus norv
11	7	4.3	216	1	O1E5_HUMAN	Q9um60 homo sapien
12	7	4.3	216	1	OL1F_HUMAN	P47886 homo sapien
13	7	4.3	216	1	OL1I_HUMAN	P47889 homo sapien
14	7	4.3	216	1	OL1L_HUMAN	P47892 homo sapien
15	7	4.3	222	1	OLF6_MOUSE	P34986 mus musculu
16	7	4.3	234	1	GU33_RAT	P35896 rattus norv
17	7	4.3	234	1	GU38_RAT	P35897 rattus norv
18	7	4.3	275	1	MYOD_BRARE	Q90477 brachydanio
19	7	4.3	308	1	O2D2_HUMAN	Q9h210 homo sapien
20	7	4.3	309	1	O1A1_HUMAN	Q9p1q5 homo sapien
21	7	4.3	311	1	O2T1_HUMAN	O43869 homo sapien
22	7	4.3	312	1	GU27_RAT	P34987 rattus norv
23	7	4.3	312	1	O1F1_HUMAN	O43749 homo sapien
24	7	4.3	314	1	O1E1_HUMAN	P30953 homo sapien
25	7	4.3	314	1	OLF9_RAT	P23272 rattus norv
26	7	4.3	315	1	O3A1_HUMAN	P47881 homo sapien
27	7	4.3	315	1	O3A2_HUMAN	P47893 homo sapien
28	7	4.3	315	1	O3A3_HUMAN	P47888 homo sapien
29	7	4.3	316	1	OCD3_HUMAN	Q9ugf7 homo sapien
30	7	4.3	318	1	OYD1_HUMAN	Q9h346 homo sapien
31	7	4.3	319	1	O2S2_HUMAN	Q9nqn1 homo sapien
32	7	4.3	319	1	OPPB_LACLA	P50989 lactococcus
33	7	4.3	320	1	O7C1_HUMAN	O76099 homo sapien
34	7	4.3	320	1	OAJ1_HUMAN	P30954 homo sapien
35	7	4.3	323	1	O1E2_HUMAN	P47887 homo sapien
36	7	4.3	330	1	RUVB_CHLMU	Q9pkz8 chlamydia m
37	7	4.3	355	1	O1I1_HUMAN	O60431 homo sapien
38	7	4.3	492	1	FSC2_BOVIN	O18728 bos taurus
39	7	4.3	492	1	FSC2_HUMAN	O14926 homo sapien
40	7	4.3	606	1	WDR1_HUMAN	O75083 homo sapien
41	7	4.3	608	1	ALBU_CANFA	P49822 canis famil
42	7	4.3	608	1	ALBU_FELCA	P49064 felis silve
43	7	4.3	608	1	ALBU_MOUSE	P07724 mus musculu
44	7	4.3	608	1	ALBU_RAT	P02770 rattus norv
45	7	4.3	609	1	ALBU_MERUN	O35090 meriones un

ALIGNMENTS

RESULT 1
OLF8_CHICK
ID OLF8_CHICK STANDARD; PRT; 159 AA.
AC Q98913;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Olfactory receptor-like protein COR8 (Fragment).
GN COR8.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=96317247; PubMed=8734500;
RA Nef S., Allaman I., Fiumelli H., de Castro E., Nef P.;
RT "Olfaction in birds: differential embryonic expression of nine
RT putative odorant receptor genes in the avian olfactory system.";
RL Mech. Dev. 55:65-77(1996).
CC -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC ----------

CC ----------
DR EMBL; Z79591; CAB01852.1; -.
DR GCRDb; GCR_1091; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Olfaction; Multigene family.
FT NON_TER 1 1
FT DOMAIN <1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 41 4 (POTENTIAL).
FT DOMAIN 42 82 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 83 103 5 (POTENTIAL).
FT DOMAIN 104 116 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 117 137 6 (POTENTIAL).
FT DOMAIN 138 148 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 149 >159 7 (POTENTIAL).
FT NON_TER 159 159
SQ SEQUENCE 159 AA; 17385 MW; B1B5619981A8B4E4 CRC64;

Query Match 5.6%; Score 9; DB 1; Length 159;

```
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
Db      1 VAICNPLLY 9


RESULT   2
OL7B_MOUSE
ID   OL7B_MOUSE     STANDARD;      PRT;   222 AA.
AC   P34983;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor 7B (K4) (Fragment).
GN   OLFR7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=93258822; PubMed=7683976;
RA   Ressler K.J., Sullivan S.L., Buck L.B.;
RT   "A zonal organization of odorant receptor gene expression in the
RT   olfactory epithelium.";
RL   Cell 73:597-609(1993).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: OLFACTORY RECEPTORS 7A-7I ARE PRODUCED BY EIGHT
CC       DIFFERENT GENES WITHIN THE OLFR7 COMPLEX.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14569; AAA39853.1; -.
DR   PIR; C40745; C40745.
DR   GCRDb; GCR_0716; -.
DR   MGD; MGI:104712; Olfr7.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Multigene family;
KW   Olfaction.
FT   NON_TER       1      1
FT   TRANSMEM     <1     19       2 (POTENTIAL).
FT   DOMAIN       20     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     41     60       3 (POTENTIAL).
FT   DOMAIN       61     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     98       4 (POTENTIAL).
FT   DOMAIN       99    136       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    137    159       5 (POTENTIAL).
FT   DOMAIN      160    176       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    177    200       6 (POTENTIAL).
FT   DOMAIN      201    212       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    213   >222       7 (POTENTIAL).
FT   NON_TER     222    222
SQ   SEQUENCE   222 AA;  24855 MW;  5C1978B5C93FC3E1 CRC64;


  Query Match             5.6%;  Score 9;  DB 1;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 0.088;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
Db     64 VAICNPLLY 72


RESULT   3
O8B8_HUMAN
ID   O8B8_HUMAN     STANDARD;      PRT;   311 AA.
AC   Q15620;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 8B8 (Olfactory receptor TPCR85).
GN   OR8B8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 126-282 FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X89675; CAA61822.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     25       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     26     49       1 (POTENTIAL).
FT   DOMAIN       50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     79       2 (POTENTIAL).
FT   DOMAIN       80    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    120       3 (POTENTIAL).
FT   DOMAIN      121    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    158       4 (POTENTIAL).
FT   DOMAIN      159    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    219       5 (POTENTIAL).
FT   DOMAIN      220    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    259       6 (POTENTIAL).
FT   DOMAIN      260    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    311       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     97    189       BY SIMILARITY.
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    204    204       G -> S (IN REF. 1).
SQ   SEQUENCE   311 AA;  34482 MW;  9D8AAC4DA179ABE6 CRC64;


  Query Match             5.6%;  Score 9;  DB 1;  Length 311;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
```

```
Db    124 VAICNPLLY 132


RESULT   4
OLF1_CANFA
ID   OLF1_CANFA     STANDARD;      PRT;   311 AA.
AC   Q95154;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Olfactory receptor-like protein OLF1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97008103; PubMed=8855279;
RA   Issel-Tarver L., Rine J.;
RT   "Organization and expression of canine olfactory receptor genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:10897-10902(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT OR SPERM CELL RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U53679; AAB37239.1; -.
DR   GCRDb; GCR_1192; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     24       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     25     48       1 (POTENTIAL).
FT   DOMAIN       49     56       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     57     78       2 (POTENTIAL).
FT   DOMAIN       79     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    100    119       3 (POTENTIAL).
FT   DOMAIN      120    138       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    139    157       4 (POTENTIAL).
FT   DOMAIN      158    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    218       5 (POTENTIAL).
FT   DOMAIN      219    235       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    236    259       6 (POTENTIAL).
FT   DOMAIN      260    271       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    291       7 (POTENTIAL).
FT   DOMAIN      292    311       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   311 AA;  35209 MW;  05A2132474F543A1 CRC64;


  Query Match             5.6%;  Score 9;  DB 1;  Length 311;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
Db    123 VAICNPLLY 131


RESULT   5
O5I1_HUMAN
ID   O5I1_HUMAN     STANDARD;      PRT;   314 AA.
AC   Q13606;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 5I1 (Olfactory receptor-like protein OLF1).
GN   OR5I1 OR OLF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Issel-Tarver L., Rine J.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U56420; AAB01214.1; -.
DR   GCRDb; GCR_1925; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00245; OLFACTORYR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     29     51       1 (POTENTIAL).
FT   DOMAIN       52     59       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     60     81       2 (POTENTIAL).
FT   DOMAIN       82    102       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    103    122       3 (POTENTIAL).
FT   DOMAIN      123    141       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    142    166       4 (POTENTIAL).
FT   DOMAIN      167    207       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    208    228       5 (POTENTIAL).
FT   DOMAIN      229    241       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    242    262       6 (POTENTIAL).
FT   DOMAIN      263    273       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    274    294       7 (POTENTIAL).
FT   DOMAIN      295    314       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    191       BY SIMILARITY.
FT   CARBOHYD      7      7       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   314 AA;  36048 MW;  334BE9BF3D9A4D63 CRC64;


  Query Match             5.6%;  Score 9;  DB 1;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
Db    126 VAICNPLLY 134


RESULT   6
VE5_HPV35
ID   VE5_HPV35      STANDARD;      PRT;    81 AA.
AC   P27226;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Probable E5 protein.
GN   E5.
```

```
OS   Human papillomavirus type 35.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10587;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92124753; PubMed=1310198;
RA   Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT   "The phylogenetic relationship and complete nucleotide sequence of
RT   human papillomavirus type 35.";
RL   Virology 186:770-776(1992).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74117; AAA46970.1; -.
DR   PIR; D40824; W5WL35.
DR   InterPro; IPR004270; Papilloma_E5.
DR   Pfam; PF03025; Papilloma_E5; 1.
KW   Early protein.
SQ   SEQUENCE   81 AA;  9000 MW;  B12A23102E72163B CRC64;


  Query Match             4.3%;  Score 7;  DB 1;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 4.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    122 LLSVSLY 128
          |||||||
Db     32 LLSVSLY 38


RESULT   7
OL7H_MOUSE
ID   OL7H_MOUSE     STANDARD;      PRT;   111 AA.
AC   Q60893;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor 7H (M71) (Fragment).
GN   OLFR7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=96149403; PubMed=8570653;
RA   Sullivan S.L., Adamson M.C., Ressler K.J., Kozak C.A., Buck L.B.;
RT   "The chromosomal distribution of mouse odorant receptor genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:884-888(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: OLFACTORY RECEPTORS 7A-7I ARE PRODUCED BY EIGHT
CC       DIFFERENT GENES WITHIN THE OLFR7 COMPLEX.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28782; AAC52405.1; -.
DR   GCRDb; GCR_1678; -.
DR   MGD; MGI:104712; Olfr7.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Multigene family;
KW   Olfaction.
FT   NON_TER       1      1
FT   DOMAIN       <1     12       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     13     33       4 (POTENTIAL).
FT   DOMAIN       34     72       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     73     94       5 (POTENTIAL).
FT   DOMAIN       95    108       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    109   >111       6 (POTENTIAL).
FT   NON_TER     111    111
SQ   SEQUENCE   111 AA;  12118 MW;  64D3ADFD4BB877D3 CRC64;


  Query Match             4.3%;  Score 7;  DB 1;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 5.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFS 116
          |||||||
Db    105 GRSKAFS 111


RESULT   8
OL7C_MOUSE
ID   OL7C_MOUSE     STANDARD;      PRT;   112 AA.
AC   Q60882;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor 7C (K21) (Fragment).
GN   OLFR7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=96149403; PubMed=8570653;
RA   Sullivan S.L., Adamson M.C., Ressler K.J., Kozak C.A., Buck L.B.;
RT   "The chromosomal distribution of mouse odorant receptor genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:884-888(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: OLFACTORY RECEPTORS 7A-7I ARE PRODUCED BY EIGHT
CC       DIFFERENT GENES WITHIN THE OLFR7 COMPLEX.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28771; AAC52394.1; -.
DR   GCRDb; GCR_1690; -.
DR   MGD; MGI:104712; Olfr7.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Multigene family;
KW   Olfaction.
FT   NON_TER       1      1
FT   DOMAIN       <1     12       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     13     33       4 (POTENTIAL).
```

```
FT   DOMAIN       34     73       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     74     95       5 (POTENTIAL).
FT   DOMAIN       96    109       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    110   >112       6 (POTENTIAL).
FT   NON_TER     112    112
SQ   SEQUENCE   112 AA;  12261 MW;  8C1AD5ABBA6C78BB CRC64;


  Query Match             4.3%;  Score 7;  DB 1;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GRSKAFS 116
           |||||||
Db     106 GRSKAFS 112


RESULT   9
OL7D_MOUSE
ID   OL7D_MOUSE     STANDARD;      PRT;   112 AA.
AC   Q60884;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor 7D (M15) (Fragment).
GN   OLFR7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=96149403; PubMed=8570653;
RA   Sullivan S.L., Adamson M.C., Ressler K.J., Kozak C.A., Buck L.B.;
RT   "The chromosomal distribution of mouse odorant receptor genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:884-888(1996).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: OLFACTORY RECEPTORS 7A-7I ARE PRODUCED BY EIGHT
CC       DIFFERENT GENES WITHIN THE OLFR7 COMPLEX.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28773; AAC52396.1; -.
DR   GCRDb; GCR_1688; -.
DR   MGD; MGI:104712; Olfr7.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Multigene family;
KW   Olfaction.
FT   NON_TER       1      1
FT   DOMAIN       <1     12       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     13     33       4 (POTENTIAL).
FT   DOMAIN       34     73       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     74     95       5 (POTENTIAL).
FT   DOMAIN       96    109       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    110   >112       6 (POTENTIAL).
FT   NON_TER     112    112
SQ   SEQUENCE   112 AA;  12426 MW;  08640539B168518E CRC64;


  Query Match             4.3%;  Score 7;  DB 1;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GRSKAFS 116
           |||||||
Db     106 GRSKAFS 112


RESULT  10
GU01_RAT
ID   GU01_RAT       STANDARD;      PRT;   185 AA.
AC   P35894;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Possible gustatory receptor clone PTE01 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FISCHER; TISSUE=Lingual epithelium;
RX   MEDLINE=93138133; PubMed=8380780;
RA   Abe K., Kusakabe Y., Tanemura K., Emori Y., Arai S.;
RT   "Multiple genes for G protein-coupled receptors and their expression
RT   in lingual epithelia.";
RL   FEBS Lett. 316:253-256(1993).
CC   -!- FUNCTION: POSSIBLE TASTE RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: TONGUE-SPECIFIC.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   PIR; S28996; S28996.
DR   GCRDb; GCR_0825; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   NON_TER       1      1
FT   TRANSMEM     <1     11       2 (POTENTIAL).
FT   DOMAIN       12     42       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     43     62       3 (POTENTIAL).
FT   DOMAIN       63     84       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     85    105       4 (POTENTIAL).
FT   DOMAIN      106    138       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    139    160       5 (POTENTIAL).
FT   DOMAIN      161    182       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    183   >185       6 (POTENTIAL).
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20605 MW;  FCF72E763D578F23 CRC64;


  Query Match             4.3%;  Score 7;  DB 1;  Length 185;
  Best Local Similarity   100.0%;  Pred. No. 8.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     178 KAFSTCG 184


RESULT  11
O1E5_HUMAN
ID   O1E5_HUMAN     STANDARD;      PRT;   216 AA.
AC   Q9UM60;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor 1E5 (Olfactory receptor 13-66) (OR13-66)
DE   (Fragment).
GN   OR1E5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98160182; PubMed=9500546;
RA   Rouquier S., Taviaux S., Trask B.J., Brand-Arpon V., van den Engh G.,
RA   Demaille J.G., Giorgi D.;
RT   "Distribution of olfactory receptor genes in the human genome.";
RL   Nat. Genet. 18:243-250(1998).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U86222; AAC39613.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   NON_TER       1      1
FT   TRANSMEM     <1     12       2 (POTENTIAL).
FT   DOMAIN       13     33       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     34     53       3 (POTENTIAL).
FT   DOMAIN       54     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     91       4 (POTENTIAL).
FT   DOMAIN       92    129       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    130    152       5 (POTENTIAL).
FT   DOMAIN      153    169       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    170    192       6 (POTENTIAL).
FT   DOMAIN      193    204       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206   >216       7 (POTENTIAL).
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23977 MW;  0A48A77D1A98C244 CRC64;

  Query Match             4.3%;  Score 7;  DB 1;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     169 KAFSTCG 175


RESULT  12
OL1F_HUMAN
ID   OL1F_HUMAN     STANDARD;      PRT;   216 AA.
AC   P47886;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor-like protein OR17-82 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94272458; PubMed=8004088;
RA   Ben-Arie N., Lancet D., Taylor C., Khen M., Walker N.,
RA   Ledbetter D.H., Carrozzo R., Patel K., Sheer D., Lehrach H.,
RA   North M.A.;
RT   "Olfactory receptor gene cluster on human chromosome 17: possible
RT   duplication of an ancestral receptor repertoire.";
RL   Hum. Mol. Genet. 3:229-235(1994).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U04684; AAA18348.1; -.
DR   GCRDb; GCR_0851; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   NON_TER       1      1
FT   TRANSMEM     <1     12       2 (POTENTIAL).
FT   DOMAIN       13     33       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     34     53       3 (POTENTIAL).
FT   DOMAIN       54     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     74     91       4 (POTENTIAL).
FT   DOMAIN       92    129       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    130    153       5 (POTENTIAL).
FT   DOMAIN      154    170       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    171    194       6 (POTENTIAL).
FT   DOMAIN      195    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206   >216       7 (POTENTIAL).
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23634 MW;  ACC757FB3D9BB55F CRC64;

  Query Match             4.3%;  Score 7;  DB 1;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     169 KAFSTCG 175


RESULT  13
OL1I_HUMAN
ID   OL1I_HUMAN     STANDARD;      PRT;   216 AA.
AC   P47889;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor-like protein OR17-207 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94272458; PubMed=8004088;
RA   Ben-Arie N., Lancet D., Taylor C., Khen M., Walker N.,
RA   Ledbetter D.H., Carrozzo R., Patel K., Sheer D., Lehrach H.,
RA   North M.A.;
RT   "Olfactory receptor gene cluster on human chromosome 17: possible
RT   duplication of an ancestral receptor repertoire.";
RL   Hum. Mol. Genet. 3:229-235(1994).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U04687; AAA18350.1; -.
DR   GCRDb; GCR_0854; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   NON_TER       1      1
FT   TRANSMEM     <1     12       2 (POTENTIAL).
FT   DOMAIN       13     33       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     34     53       3 (POTENTIAL).
FT   DOMAIN       54     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     74     91       4 (POTENTIAL).
FT   DOMAIN       92    129       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    130    153       5 (POTENTIAL).
FT   DOMAIN      154    170       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    171    194       6 (POTENTIAL).
FT   DOMAIN      195    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206   >216       7 (POTENTIAL).
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23617 MW;  28DBCBE39733F9F2 CRC64;

  Query Match              4.3%;  Score 7;  DB 1;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     169 KAFSTCG 175


RESULT  14
OL1L_HUMAN
ID   OL1L_HUMAN     STANDARD;      PRT;   216 AA.
AC   P47892;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Olfactory receptor-like protein OR17-219 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94272458; PubMed=8004088;
RA   Ben-Arie N., Lancet D., Taylor C., Khen M., Walker N.,
RA   Ledbetter D.H., Carrozzo R., Patel K., Sheer D., Lehrach H.,
RA   North M.A.;
RT   "Olfactory receptor gene cluster on human chromosome 17: possible
RT   duplication of an ancestral receptor repertoire.";
RL   Hum. Mol. Genet. 3:229-235(1994).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U04691; AAA18354.1; -.
DR   GCRDb; GCR_0858; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   NON_TER       1      1
FT   TRANSMEM     <1     12       2 (POTENTIAL).
FT   DOMAIN       13     33       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     34     53       3 (POTENTIAL).
FT   DOMAIN       54     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     74     91       4 (POTENTIAL).
FT   DOMAIN       92    129       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    130    153       5 (POTENTIAL).
FT   DOMAIN      154    170       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    171    194       6 (POTENTIAL).
FT   DOMAIN      195    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206   >216       7 (POTENTIAL).
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23613 MW;  74449CEC8C73144C CRC64;

  Query Match              4.3%;  Score 7;  DB 1;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     113 KAFSTCG 119
           |||||||
Db     169 KAFSTCG 175


RESULT  15
OLF6_MOUSE
ID   OLF6_MOUSE     STANDARD;      PRT;   222 AA.
AC   P34986;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Olfactory receptor 6 (M50) (Fragment).
GN   OLFR6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=93258822; PubMed=7683976;
RA   Ressler K.J., Sullivan S.L., Buck L.B.;
RT   "A zonal organization of odorant receptor gene expression in the
RT   olfactory epithelium.";
RL   Cell 73:597-609(1993).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: OLFACTORY EPITHELIUM.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L14567; AAA39851.1; -.
DR   PIR; B40745; B40745.
DR   GCRDb; GCR_0714; -.
DR   MGD; MGI:104713; Olfr6.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
```

```
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Multigene family;
KW   Olfaction.
FT   NON_TER       1      1
FT   TRANSMEM     <1     19       2 (POTENTIAL).
FT   DOMAIN       20     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     41     60       3 (POTENTIAL).
FT   DOMAIN       61     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     98       4 (POTENTIAL).
FT   DOMAIN       99    136       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    137    159       5 (POTENTIAL).
FT   DOMAIN      160    176       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    177    200       6 (POTENTIAL).
FT   DOMAIN      201    212       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    213   >222       7 (POTENTIAL).
FT   NON_TER     222    222
SQ   SEQUENCE   222 AA;  24748 MW;  7235CEDDDDE43BE7 CRC64;


  Query Match             4.3%;  Score 7;  DB 1;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     133 IFMYVRP 139
           |||||||
Db     196 IFMYVRP 202


Search completed: June 27, 2002, 14:17:10
Job time: 252 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 27, 2002, 14:12:23 ; Search time 23.9 Seconds
(without alignments)
1172.602 Million cell updates/sec

Title: US-09-747-155-225
Perfect score: 162
Sequence: 1 VAICNPLLYPVMMSNKLSAQ.....LAEDQDKVYSLFYTIIIPLL 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	162	100.0	162	4	Q9NZP5	Q9nzp5 homo sapien
2	74	45.7	162	6	Q9N1M1	Q9n1m1 pan troglod
3	10	6.2	162	6	Q9N1N4	Q9n1n4 papio hamad
4	10	6.2	221	11	Q9Z230	Q9z230 rattus norv
5	10	6.2	309	11	Q9EQA6	Q9eqa6 mus musculu
6	10	6.2	309	11	Q9EQA4	Q9eqa4 mus musculu
7	10	6.2	314	11	Q9EQA5	Q9eqa5 mus musculu
8	10	6.2	315	6	Q9TSM7	Q9tsm7 macaca fasc
9	9	5.6	157	4	Q15614	Q15614 homo sapien
10	9	5.6	157	4	Q15617	Q15617 homo sapien
11	9	5.6	157	6	Q28307	Q28307 canis famil
12	9	5.6	157	11	Q63595	Q63595 rattus norv
13	9	5.6	161	6	Q9N1G5	Q9n1g5 saimiri bol
14	9	5.6	162	6	Q9N1N3	Q9n1n3 papio hamad
15	9	5.6	162	6	Q9N1L5	Q9n1l5 gorilla gor
16	9	5.6	162	6	Q9N1L0	Q9n1l0 eulemur ful
17	9	5.6	162	6	Q9N1J9	Q9n1j9 macaca sylv
18	9	5.6	162	6	Q9N1J2	Q9n1j2 callithrix
19	9	5.6	162	6	Q9N1I9	Q9n1i9 callithrix
20	9	5.6	162	6	Q9N1H0	Q9n1h0 saimiri sci
21	9	5.6	162	6	Q9N1G2	Q9n1g2 saimiri bol
22	9	5.6	214	11	Q9JM36	Q9jm36 mus musculu
23	9	5.6	215	4	Q96RD0	Q96rd0 homo sapien
24	9	5.6	215	4	Q96RC9	Q96rc9 homo sapien
25	9	5.6	215	4	Q96RC6	Q96rc6 homo sapien
26	9	5.6	216	4	Q96RC8	Q96rc8 homo sapien
27	9	5.6	216	4	Q96RC2	Q96rc2 homo sapien
28	9	5.6	216	4	Q96RB5	Q96rb5 homo sapien
29	9	5.6	216	4	Q96RB4	Q96rb4 homo sapien
30	9	5.6	216	4	Q96RB3	Q96rb3 homo sapien
31	9	5.6	216	4	Q96RB2	Q96rb2 homo sapien
32	9	5.6	216	11	Q9JM27	Q9jm27 mus musculu
33	9	5.6	217	4	Q96RC7	Q96rc7 homo sapien
34	9	5.6	217	4	Q96RC3	Q96rc3 homo sapien
35	9	5.6	217	4	Q96R39	Q96r39 homo sapien
36	9	5.6	221	11	Q9Z236	Q9z236 rattus norv
37	9	5.6	221	11	Q9Z231	Q9z231 rattus norv
38	9	5.6	223	11	Q9Z1V5	Q9z1v5 mus musculu
39	9	5.6	304	11	Q9QW36	Q9qw36 rattus sp.
40	9	5.6	306	11	Q9EQ89	Q9eq89 mus musculu
41	9	5.6	307	11	Q9EQ96	Q9eq96 mus musculu
42	9	5.6	309	11	Q9EQG1	Q9eqg1 mus musculu
43	9	5.6	309	11	Q9ERU6	Q9eru6 mus musculu
44	9	5.6	310	11	Q9EQB0	Q9eqb0 mus musculu
45	9	5.6	310	11	Q9EQA7	Q9eqa7 mus musculu

ALIGNMENTS

RESULT 1
Q9NZP5
ID Q9NZP5 PRELIMINARY; PRT; 162 AA.
AC Q9NZP5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE OLFACTORY RECEPTOR (FRAGMENT).
GN HSA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20183981; PubMed=10706615;
RA Rouquier S., Blancher A., Giorgi D.;
RT "The olfactory receptor gene repertoire in primates and mouse:
RT Evidence for reduction of the functional fraction in primates.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR EMBL; AF179759; AAF40348.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 162 162
SQ SEQUENCE 162 AA; 18386 MW; 6716C48C3ED0E4D8 CRC64;

Query Match 100.0%; Score 162; DB 4; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.9e-162;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60

Qy     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDILKKKSEKGRSKAFSTCGA 120

Qy    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162
          |||||||||||||||||||||||||||||||||||||||||||
Db    121 HLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIPLL 162


RESULT   2
Q9N1M1
ID   Q9N1M1      PRELIMINARY;      PRT;   162 AA.
AC   Q9N1M1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   PTR210.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179735; AAF40330.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     162    162
SQ   SEQUENCE   162 AA;  18389 MW;  586280CC7BD5F4CE CRC64;


  Query Match             45.7%;  Score 74;  DB 6;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-69;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLTFCRFNIIHYFYCEILQ 60

Qy     61 LFKISCNGPSINAL 74
          ||||||||||||||
Db     61 LFKISCNGPSINAL 74


RESULT   3
Q9N1N4
ID   Q9N1N4      PRELIMINARY;      PRT;   162 AA.
AC   Q9N1N4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   PPA138.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179721; AAF40317.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     162    162
SQ   SEQUENCE   162 AA;  18112 MW;  ACD9BA1ECBD37959 CRC64;


  Query Match             6.2%;  Score 10;  DB 6;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    113 KAFSTCGAHL 122
          ||||||||||
Db    113 KAFSTCGAHL 122


RESULT   4
Q9Z230
ID   Q9Z230      PRELIMINARY;      PRT;   221 AA.
AC   Q9Z230;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HFL-VN1;
RA   Singer M.S., Hughes T.E., Shepherd G.M., Greer C.A.;
RT   "Isolation of olfactory receptor mRNA sequences from olfactory bulb
RT   glomerular layer.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF091576; AAC64596.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     221    221
SQ   SEQUENCE   221 AA;  24719 MW;  1FF1BF090F84E4C1 CRC64;


  Query Match             6.2%;  Score 10;  DB 11;  Length 221;
  Best Local Similarity   100.0%;  Pred. No. 0.03;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    109 KGRSKAFSTC 118
          ||||||||||
Db    169 KGRSKAFSTC 178


RESULT   5
Q9EQA6
ID   Q9EQA6      PRELIMINARY;      PRT;   309 AA.
AC   Q9EQA6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K25.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
```

```
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282283; AAG39868.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   309 AA;  34645 MW;  A07D063336AEF837 CRC64;


  Query Match             6.2%;  Score 10;  DB 11;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 0.04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTCG 119
          ||||||||||
Db    232 GRSKAFSTCG 241


RESULT   6
Q9EQA4
ID   Q9EQA4      PRELIMINARY;      PRT;   309 AA.
AC   Q9EQA4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K27.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282285; AAG39870.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   309 AA;  34592 MW;  D58DA04AFFDE88DD CRC64;


  Query Match             6.2%;  Score 10;  DB 11;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 0.04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTCG 119
          ||||||||||
Db    232 GRSKAFSTCG 241


RESULT   7
Q9EQA5
ID   Q9EQA5      PRELIMINARY;      PRT;   314 AA.
AC   Q9EQA5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ODORANT RECEPTOR K26.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE=21015403; PubMed=11130974;
RA   Xie S.Y., Feinstein P., Mombaerts P.;
RT   "Characterization of a Cluster Comprising 100 Odorant Receptor Genes
RT   in Mouse.";
RL   Mamm. Genome 11:1070-1078(2000).
DR   EMBL; AF282284; AAG39869.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   314 AA;  35120 MW;  4D1C19E3ABEED439 CRC64;


  Query Match             6.2%;  Score 10;  DB 11;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 0.041;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTCG 119
          ||||||||||
Db    237 GRSKAFSTCG 246


RESULT   8
Q9TSM7
ID   Q9TSM7      PRELIMINARY;      PRT;   315 AA.
AC   Q9TSM7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MFOR1P1;
RX   MEDLINE=99443869; PubMed=10512677;
RA   Sharon D., Glusman G., Pilpel Y., Khen M., Gruetzner F., Haaf T.,
RA   Lancet D.;
RT   "Primate evolution of an olfactory receptor cluster: diversification
RT   by gene conversion and recent emergence of pseudogenes.";
RL   Genomics 61:24-36(1999).
DR   EMBL; AF101778; AAF03348.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   315 AA;  34638 MW;  778E19FBA03F646C CRC64;


  Query Match             6.2%;  Score 10;  DB 6;  Length 315;
  Best Local Similarity   100.0%;  Pred. No. 0.041;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    113 KAFSTCGAHL 122
          ||||||||||
Db    239 KAFSTCGAHL 248


RESULT   9
Q15614
ID   Q15614      PRELIMINARY;      PRT;   157 AA.
AC   Q15614;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   PUTATIVE OLFACTORY RECEPTOR (FRAGMENT).
GN   TPCR120.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
DR   EMBL; X89669; CAA61816.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     157    157
SQ   SEQUENCE   157 AA;  17300 MW;  64B3DD9A484D1E55 CRC64;


  Query Match             5.6%;  Score 9;  DB 4;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTC 118
          |||||||||
Db    108 GRSKAFSTC 116


RESULT  10
Q15617
ID   Q15617      PRELIMINARY;      PRT;   157 AA.
AC   Q15617;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE OLFACTORY RECEPTOR (FRAGMENT).
GN   TPCR25.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Fuchs T., Malecova B., Linhart C., Sharan R., Khen M., Herwig R.,
RA   Shmulevich D., Elkon R., Steinfath M., O'Brien J.K., Radelof U.,
RA   Lehrach H., Olender Z., Gluesman G., Lancet D., Shamir R.;
RT   "DEFOG: A Practical Scheme for Deciphering Families of Genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X89672; CAA61819.1; -.
DR   EMBL; AF399507; AAK94992.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     157    157
SQ   SEQUENCE   157 AA;  17264 MW;  0F73893D7D2A2E04 CRC64;


  Query Match             5.6%;  Score 9;  DB 4;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTC 118
          |||||||||
Db    108 GRSKAFSTC 116


RESULT  11
Q28307
ID   Q28307      PRELIMINARY;      PRT;   157 AA.
AC   Q28307;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE OLFACTORY RECEPTOR (FRAGMENT).
GN   TPCR79.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
DR   EMBL; X89665; CAA61812.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     157    157
SQ   SEQUENCE   157 AA;  17236 MW;  BEC8C82DB37A6AAD CRC64;


  Query Match             5.6%;  Score 9;  DB 6;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTC 118
          |||||||||
Db    108 GRSKAFSTC 116


RESULT  12
Q63595
ID   Q63595      PRELIMINARY;      PRT;   157 AA.
AC   Q63595;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE OLFACTORY RECEPTOR (FRAGMENT).
GN   TPCR21.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97224452; PubMed=9119360;
RA   Vanderhaeghen P., Schurmans S., Vassart G., Parmentier M.;
RT   "Specific repertoire of olfactory receptor genes in the male germ
RT   cells of several mammalian species.";
RL   Genomics 39:239-246(1997).
DR   EMBL; X89704; CAA61851.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
```

```
FT   NON_TER       1      1
FT   NON_TER     157    157
SQ   SEQUENCE   157 AA;  17653 MW;  84724899D60CB049 CRC64;


  Query Match             5.6%;  Score 9;  DB 11;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTC 118
          |||||||||
Db    108 GRSKAFSTC 116


RESULT  13
Q9N1G5
ID   Q9N1G5      PRELIMINARY;      PRT;   161 AA.
AC   Q9N1G5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   SBO218.
OS   Saimiri boliviensis (Bolivian squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=27679;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179838; AAF40404.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     161    161
SQ   SEQUENCE   161 AA;  17370 MW;  446162DD51F73E81 CRC64;


  Query Match             5.6%;  Score 9;  DB 6;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
Db      1 VAICNPLLY 9


RESULT  14
Q9N1N3
ID   Q9N1N3      PRELIMINARY;      PRT;   162 AA.
AC   Q9N1N3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   PPA139.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179722; AAF40318.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     162    162
SQ   SEQUENCE   162 AA;  18082 MW;  44E77F9E7E4373D4 CRC64;


  Query Match             5.6%;  Score 9;  DB 6;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VAICNPLLY 9
          |||||||||
Db      1 VAICNPLLY 9


RESULT  15
Q9N1L5
ID   Q9N1L5      PRELIMINARY;      PRT;   162 AA.
AC   Q9N1L5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OLFACTORY RECEPTOR (FRAGMENT).
GN   GGO107.
OS   Gorilla gorilla (gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX   NCBI_TaxID=9593;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20183981; PubMed=10706615;
RA   Rouquier S., Blancher A., Giorgi D.;
RT   "The olfactory receptor gene repertoire in primates and mouse:
RT   Evidence for reduction of the functional fraction in primates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
DR   EMBL; AF179756; AAF40346.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     162    162
SQ   SEQUENCE   162 AA;  17360 MW;  76324350D05E53E6 CRC64;


  Query Match             5.6%;  Score 9;  DB 6;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    110 GRSKAFSTC 118
          |||||||||
Db    110 GRSKAFSTC 118


Search completed: June 27, 2002, 14:16:38
Job time: 255 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 11:17:07 ; Search time 1624.99 Seconds
(without alignments)
4044.959 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	21	4.3	389	10	BM253378	BM253378 514417 MA
c	2	21	4.3	496	10	BE752648	BE752648 205097 MA
	3	20	4.1	281	9	BB579415	BB579415 BB579415
	4	20	4.1	408	10	BE991812	BE991812 UI-M-BZ1-
	5	20	4.1	669	12	AG102096	AG102096 Pan trogl
c	6	20	4.1	811	12	AQ270566	AQ270566 HS_2046_B
c	7	19	3.9	279	9	AA689627	AA689627 vs08h11.r
	8	19	3.9	300	9	AV176153	AV176153 AV176153
c	9	19	3.9	320	10	BE523391	BE523391 M36D6STM
c	10	19	3.9	327	9	AA611944	AA611944 vo86c02.r
	11	19	3.9	348	9	BE148808	BE148808 CM4-HT024
	12	19	3.9	355	10	BG733743	BG733743 As_nc_01H
c	13	19	3.9	371	12	AQ241320	AQ241320 CITBI-E1-
	14	19	3.9	378	10	M79761	M79761 wEST00298 M
c	15	19	3.9	408	9	AW835801	AW835801 QV4-LT001
c	16	19	3.9	408	12	AQ593534	AQ593534 HS_5445_B
c	17	19	3.9	410	9	AW835811	AW835811 QV4-LT001
c	18	19	3.9	410	12	AQ674479	AQ674479 HS_5478_B
	19	19	3.9	423	9	AW862208	AW862208 QV4-CT036
	20	19	3.9	447	9	AW453392	AW453392 660033A04
c	21	19	3.9	447	10	BI034798	BI034798 RC4-NN117
c	22	19	3.9	452	10	BI497566	BI497566 sag24c11.
c	23	19	3.9	467	12	BH298263	BH298263 CH230-146
	24	19	3.9	477	9	AW871068	AW871068 ra54d03.y
c	25	19	3.9	496	12	AZ366926	AZ366926 1M0116D24
	26	19	3.9	502	9	AI605706	AI605706 ma60h03.y
c	27	19	3.9	513	9	AV526912	AV526912 AV526912
	28	19	3.9	528	12	BH008916	BH008916 ef21f04.x
c	29	19	3.9	542	9	AW941540	AW941540 LD04013.3
	30	19	3.9	566	12	AZ874840	AZ874840 2M0189B10
	31	19	3.9	607	12	AQ548864	AQ548864 RPCI-11-3
c	32	19	3.9	629	9	AW384082	AW384082 RC0-HT037
c	33	19	3.9	641	12	BH291135	BH291135 CH230-78D
c	34	19	3.9	655	12	AZ086983	AZ086983 RPCI-23-6
c	35	19	3.9	655	12	AG151658	AG151658 Pan trogl
	36	19	3.9	698	10	BJ134615	BJ134615 BJ134615
	37	19	3.9	714	10	BE618995	BE618995 601472804
c	38	19	3.9	933	12	CNS04ZKQ	AL314387 Tetraodon
	39	19	3.9	957	12	CNS05J69	AL339786 Tetraodon
	40	19	3.9	964	12	AQ899367	AQ899367 HS_2057_A
c	41	18	3.7	150	10	BI052471	BI052471 PM0-GN018
	42	18	3.7	158	9	BE014092	BE014092 125674 MA
c	43	18	3.7	162	12	AZ113993	AZ113993 RPCI-23-4
	44	18	3.7	179	10	M79844	M79844 wEST00381 M
c	45	18	3.7	205	9	BE172103	BE172103 MR0-HT055

ALIGNMENTS

RESULT 1
BM253378/c
LOCUS BM253378 389 bp mRNA linear EST 17-DEC-2001
DEFINITION 514417 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM253378
VERSION BM253378.1 GI:17888977
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 389)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 110 row: F column: 4
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .389
/organism="Bos taurus"
/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendonosus muscle, and fetal longissimus muscle."
BASE COUNT 130 a 51 c 61 g 147 t
ORIGIN

Query Match 4.3%; Score 21; DB 10; Length 389;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 tttatttttggtgcttttata 250
||||||||||||||||||||
Db 64 TTTATTTTTGGTGCTTTTATA 44

RESULT 2
BE752648/c
LOCUS BE752648 496 bp mRNA linear EST 25-APR-2001
DEFINITION 205097 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE752648
VERSION BE752648.1 GI:10166640
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 496)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 53 row: F column: 16
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .496
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from testis, thymus, semitendonosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."
BASE COUNT 171 a 66 c 81 g 178 t
ORIGIN

Query Match 4.3%; Score 21; DB 10; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 tttatttttggtgcttttata 250
||||||||||||||||||||
Db 156 TTTATTTTTGGTGCTTTTATA 136

RESULT 3
BB579415
LOCUS BB579415 281 bp mRNA linear EST 30-NOV-2000
DEFINITION BB579415 RIKEN full-length enriched, 11 days embryo gonad Mus musculus cDNA clone 7030401H04 5', mRNA sequence.
ACCESSION BB579415
VERSION BB579415.1 GI:11475362
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 281)
AUTHORS Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodoyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .281
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7030401H04"
/clone_lib="RIKEN full-length enriched, 11 days embryo gonad"
/sex="mixed"
/tissue_type="gonad"
/dev_stage="11 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 93.0 Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT 93 a 61 c 64 g 62 t 1 others
ORIGIN

Query Match 4.1%; Score 20; DB 9; Length 281;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 306 ttctgaaaaaaaagtctgaa 325
||||||||||||||||||||
Db 165 TTCTGAAAAAAAAGTCTGAA 184

RESULT 4
BE991812
LOCUS BE991812 408 bp mRNA linear EST 05-OCT-2000
DEFINITION UI-M-BZ1-bdz-g-01-0-UI.s1 NIH_BMAP_MHI2_S1 Mus musculus cDNA clone UI-M-BZ1-bdz-g-01-0-UI 3', mRNA sequence.
ACCESSION BE991812
VERSION BE991812.1 GI:10672004
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 408)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements The following repetitive elements were found in this cDNA sequence: 237-373, >B1_MM#SINE/Alu
Seq primer: M13 Forward
POLYA=No.
FEATURES Location/Qualifiers
source 1. .408
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BZ1-bdz-g-01-0-UI"
/clone_lib="NIH_BMAP_MHI2_S1"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_MHI2_S1 library is a subtracted library derived from NIH_BMAP_MHI2. NIH_BMAP_MHI2 is a library derived from mouse hippocampus tissue. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"
BASE COUNT 113 a 82 c 92 g 121 t
ORIGIN

Query Match 4.1%; Score 20; DB 10; Length 408;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 atattctgaaaaaaaagtct 322
||||||||||||||||||||
Db 377 ATATTCTGAAAAAAAAGTCT 396

RESULT 5
AG102096
LOCUS AG102096 669 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-105E20.R, genomic survey sequence.
ACCESSION AG102096
VERSION AG102096.1 GI:16722613
KEYWORDS GSS; GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC Library clone:PTB-105E20.R.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (sites)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 669)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES Location/Qualifiers
source 1. .669
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-105E20.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 243 a 93 c 164 g 167 t 2 others
ORIGIN

Query Match 4.1%; Score 20; DB 12; Length 669;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 308 ctgaaaaaaaagtctgaaaa 327
||||||||||||||||||||
Db 75 CTGAAAAAAAAGTCTGAAAA 94

```
RESULT   6
AQ270566/c
LOCUS       AQ270566                811 bp    DNA     linear   GSS 03-NOV-1998
DEFINITION  HS_2046_B1_F06_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2046 Col=11 Row=L, DNA sequence.
ACCESSION   AQ270566
VERSION     AQ270566.1  GI:3823161
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2046  row: L  column: 11
            Class: BAC ends
            High quality sequence stop: 811.
FEATURES             Location/Qualifiers
     source          1. .811
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2046 Col=11 Row=L"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      236 a    178 c    132 g    264 t      1 others
ORIGIN


  Query Match             4.1%;  Score 20;  DB 12;  Length 811;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    165 actgtgaaatttacaactg 184
          ||||||||||||||||||
Db    523 ACTGTGAAATTTTACAACTG 504


RESULT   7
AA689627/c
LOCUS       AA689627                279 bp    mRNA    linear   EST 16-DEC-1997
DEFINITION  vs08h11.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1137669 5', mRNA sequence.
ACCESSION   AA689627
VERSION     AA689627.1  GI:2690563
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:618941
            Putative full length read
            vector to vector length is 280
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .279
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1137669"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTT
                     T 3'];  double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead.  "
BASE COUNT       91 a     44 c     49 g     95 t
ORIGIN


  Query Match             3.9%;  Score 19;  DB 9;  Length 279;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    469 gattataattcccctgcta 487
          |||||||||||||||||||
Db     92 GATTATAATTCCCCTGCTA 74


RESULT   8
AV176153
LOCUS       AV176153                300 bp    mRNA    linear   EST 21-JUL-1999
DEFINITION  AV176153 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite
            embryo Caenorhabditis elegans cDNA clone yk508f9 3', mRNA sequence.
ACCESSION   AV176153
VERSION     AV176153.1  GI:5556054
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H.,
            Nishigaki,A., Motohashi,T., Zeng,Q., Watanabe,H., Sugimoto,A., Sano
            ,M., Miyata,A., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and
            Nomoto,H.
  TITLE     Expressed genes in C.elegans
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .300
```

```
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone="yk508f9"
                     /clone_lib="Yuji Kohara unpublished cDNA:Strain N2
                     hermaphrodite embryo"
                     /sex="hermaphrodite"
                     /dev_stage="embryo"
BASE COUNT      144 a     23 c     52 g     71 t     10 others
ORIGIN


  Query Match             3.9%;  Score 19;  DB 9;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    302 gatattctgaaaaaaaagt 320
          ||||||||||||||||||||
Db     11 GATATTCTGAAAAAAAAGT 29


RESULT  9
BE523391/c
LOCUS       BE523391                320 bp    mRNA    linear   EST 19-MAR-2001
DEFINITION  M36D6STM Arabidopsis developing seed Arabidopsis thaliana cDNA
            clone M36D6 5', mRNA sequence.
ACCESSION   BE523391
VERSION     BE523391.1  GI:9781369
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
            Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
  TITLE     A new set of Arabidopsis expressed sequence tags from developing
            seeds. The metabolic pathway from carbohydrates to seed oil
  JOURNAL   Plant Physiol. 124 (4), 1582-1594 (2000)
  MEDLINE   20567808
COMMENT     Contact: Benning, C
            Dept. of Biochemistry & Molecular Biology
            Michigan State University
            224 Biochemistry, Michigan State University, East Lansing, MI 48824
            , USA
            Tel: 517 355 1609
            Fax: 517 353 9334
            Email: benning@msu.edu
            Michigan State University DNA Sequencing Facility Arabidopsis
            Biological Resource Center, The Ohio State University, 309 Botany &
            Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
            6142920603 TEL: 6142929371.
FEATURES              Location/Qualifiers
     source           1. .320
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia"
                      /db_xref="taxon:3702"
                      /clone="M36D6"
                      /clone_lib="Arabidopsis developing seed"
                      /tissue_type="seed"
                      /dev_stage="5-13 days after flowering"
                      /lab_host="E.coli"
                      /note="Organ: Developing seed; Vector: pBluescript SK-;
                      Site_1: EcoRI; Site_2: XhoII"
BASE COUNT       86 a     53 c     88 g     92 t      1 others
ORIGIN


  Query Match             3.9%;  Score 19;  DB 10;  Length 320;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 catatgtaatcccttgctt 25
          ||||||||||||||||||
Db    195 CATATGTAATCCCTTGCTT 177


RESULT  10
AA611944/c
LOCUS       AA611944                327 bp    mRNA    linear   EST 01-OCT-1997
DEFINITION  vo86c02.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1065986 5', mRNA sequence.
ACCESSION   AA611944
VERSION     AA611944.1  GI:2461979
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 327)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:588346
            Putative full length read
            vector to vector length is 330
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 297.
FEATURES              Location/Qualifiers
     source           1. .327
                      /organism="Mus musculus"
                      /strain="FVB/N"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1065986"
                      /clone_lib="Barstead mouse irradiated colon MPLRB7"
                      /dev_stage="8 weeks"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                      from 8 week old mouse.  Colon was harvested 72 hours after
                      irradiation with 1400 Gys.  1st strand cDNA was primed
                      with  a Not I - oligo(dT) primer
                      [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT
                      T 3'];  double-stranded cDNA was ligated to Eco RI
                      adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                      into the Not I and Eco RI sites of  the modified pT7T3
                      vector. Library constructed by Bob Barstead.  "
BASE COUNT      105 a     51 c     57 g    114 t
ORIGIN


  Query Match             3.9%;  Score 19;  DB 9;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 99;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    469 gattataattcccctgcta 487
          ||||||||||||||||||||
Db    138 GATTATAATTCCCCTGCTA 120


RESULT  11
BE148808
```

```
LOCUS       BE148808                348 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  CM4-HT0243-081199-037-a06 HT0243 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE148808
VERSION     BE148808.1  GI:8611532
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 348)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0243-081
            199-037-a06&t3=1999-11-08&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 30
            High quality sequence stop: 348.
FEATURES             Location/Qualifiers
     source          1. .348
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0243"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       76 a     60 c     77 g    135 t
ORIGIN


  Query Match             3.9%;  Score 19;  DB 9;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      73 ttcatatgtaattggtttc 91
           ||||||||||||||||||
Db     301 TTCATATGTAATTGGTTTC 319


RESULT  12
BG733743
LOCUS       BG733743                355 bp    mRNA    linear   EST 11-MAY-2001
DEFINITION  As_nc_01H01_SKPL Ascaris suum (parasitic nematode) adult nerve cord
            and muscle Ascaris suum cDNA clone As_nc_01H01 5' similar to
            ref|NP_006951.1|ND3_10020 NADH dehydrogenase subunit 3 - Ascaris
            suum, mRNA sequence.
ACCESSION   BG733743
VERSION     BG733743.1  GI:14020027
KEYWORDS    EST.
SOURCE      pig roundworm.
  ORGANISM  Ascaris suum
            Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
            ; Ascarididae; Ascaris.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D., Hall
            ,N., Quayle,M. and Barrell,B.
  TITLE     Edinburgh University/Sanger Centre Nematode EST Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK.
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            The library was prepared for Dr Tim Geary, Upjohn Laboratories
            ,Kalamazoo,MI,USA Sequencing was performed by Claire Whitton ICAPB,
            Edinburgh The sequence contained a PolyA tail (trimmed)
            PCR PRimers
            FORWARD: T3
            BACKWARD: T7PL
            Plate: 01  row: H  column: 01
            Seq primer: SKPL
            High quality sequence stop: 355.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Ascaris suum"
                     /db_xref="taxon:6253"
                     /clone="As_nc_01H01"
                     /clone_lib="Ascaris suum (parasitic nematode) adult nerve
                     cord and muscle"
                     /sex="mixed"
                     /tissue_type="nerve cord and associated muscle"
                     /dev_stage="adult"
                     /note="Vector: Lambda Zap II; Site_1: EcoRI (5'end);
                     Site_2: EcoRI (3'end); Ascaris suum is an intestinal
                     nematode parasite of pigs. The library was constructed
                     from dissected nerve cord and associated muscle tissue for
                     Dr. T. Geary, Pharmacia-Upjohn Inc, Kalamazoo, MI, USA
                     [tggeary@am.pnu.com]"
BASE COUNT       68 a     15 c     78 g    194 t
ORIGIN


  Query Match             3.9%;  Score 19;  DB 10;  Length 355;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     224 ataatatttatttttggtg 242
           ||||||||||||||||||||
Db     292 ATAATATTTATTTTTGGTG 310


RESULT  13
AQ241320/c
LOCUS       AQ241320                371 bp    DNA     linear   GSS 30-SEP-1998
DEFINITION  CITBI-E1-2501016.TF.1 CITBI-E1 Homo sapiens genomic clone 2501016,
            DNA sequence.
ACCESSION   AQ241320
VERSION     AQ241320.1  GI:3673153
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 371)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CITBI-E1-2501016.TR.1
```

```
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .371
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="250l016"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
BASE COUNT       96 a     76 c    103 g     96 t
ORIGIN


  Query Match             3.9%;  Score 19;  DB 12;  Length 371;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 catgtgagtttactattgc 125
           ||||||||||||||||||
Db     294 CATGTGAGTTTACTATTGC 276


RESULT  14
M79761
LOCUS       M79761                   378 bp    mRNA    linear   EST 30-JUN-1992
DEFINITION  wEST00298 Mixed stage, Stratagene (cat. #937006) Caenorhabditis
            elegans cDNA clone CEMSD32, mRNA sequence.
ACCESSION   M79761
VERSION     M79761.1  GI:271780
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   McCombie,W.R., Adams,M.D., Kelley,J.M., FitzGerald,M.G., Utterback
            ,T.R., Khan,M., Dubnick,M., Kerlavage,A.R., Venter,J. and Fields,C.
  TITLE     Caenorhabditis elegans expressed sequence tags identify gene
            families and potential disease gene homologues
  JOURNAL   Nature Genet. 1, 124-131 (1992)
  MEDLINE   93250983
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            Seq primer: M13 Foward.
FEATURES             Location/Qualifiers
     source          1. .378
                     /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
                     /clone="CEMSD32"
                     /clone_lib="Mixed stage, Stratagene (cat. #937006)"
                     /note="Vector: Uni-ZAP XR; C. elegans mixed stage cDNA
                     library. Strategene catalog #937006. The library is oligo
                     dT primed and directionally cloned in the Uni-ZAP XR
                     vector."
BASE COUNT      163 a     32 c     78 g     98 t      7 others
ORIGIN


  Query Match             3.9%;  Score 19;  DB 10;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     302 gatattctgaaaaaaaagt 320
           ||||||||||||||||||||
Db      43 GATATTCTGAAAAAAAAGT 61


RESULT  15
AW835801/c
LOCUS       AW835801                 408 bp    mRNA    linear   EST 18-MAY-2000
DEFINITION  QV4-LT0016-240200-110-b12 LT0016 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW835801
VERSION     AW835801.1  GI:7929775
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV4-LT0016-240
            200-110-b12&t3=2000-02-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 408.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="LT0016"
                     /dev_stage="Adult"
                     /note="Organ: leiomios; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      142 a     68 c     95 g    103 t
ORIGIN


  Query Match             3.9%;  Score 19;  DB 9;  Length 408;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     360 cccatctgctttctgtctc 378
           ||||||||||||||||||||
Db     342 CCCATCTGCTTTCTGTCTC 324
```

Search completed: June 27, 2002, 12:37:17
Job time: 4810 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 10:18:27 ; Search time 1791.3 Seconds
(without alignments)
5689.289 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	487	100.0	487	6	AX181436	AX181436 Sequence
2	487	100.0	487	9	AF179759	AF179759 Homo sapi
3	486	99.8	486	6	AX242110	AX242110 Sequence
4	477.4	98.0	487	6	AX181394	AX181394 Sequence
5	477.4	98.0	487	9	AF179735	AF179735 Pan trogl
6	465	95.5	485	6	AX181383	AX181383 Sequence
7	465	95.5	485	9	AF179729	AF179729 Pan trogl
8	319	65.5	988	6	AX241993	AX241993 Sequence
9	319	65.5	150083	2	AC024892	AC024892 Homo sapi
c 10	294	60.4	138556	2	AC105564	AC105564 Rattus no
11	291.8	59.9	693	10	AY074239	AY074239 Mus muscu
12	290.4	59.6	663	10	AF091575	AF091575 Rattus no
13	282.2	57.9	918	10	AY073096	AY073096 Mus muscu
14	282.2	57.9	921	10	AY073611	AY073611 Mus muscu
15	282.2	57.9	138556	2	AC105564	AC105564 Rattus no
16	281.2	57.7	927	10	AY073094	AY073094 Mus muscu
17	279	57.3	919	10	AY074074	AY074074 Mus muscu
18	279	57.3	921	10	AY073097	AY073097 Mus muscu
19	276.4	56.8	924	10	AY073093	AY073093 Mus muscu
20	266.2	54.7	537	10	AY074165	AY074165 Mus muscu
21	260.4	53.5	918	10	AY073708	AY073708 Mus muscu
22	255	52.4	930	10	AY073612	AY073612 Mus muscu
23	250.8	51.5	930	10	AY073315	AY073315 Mus muscu
24	250.4	51.4	916	10	AY074143	AY074143 Mus muscu
25	245.4	50.4	487	6	AX181367	AX181367 Sequence
26	245.4	50.4	487	9	AF179721	AF179721 Papio ham
27	242.2	49.7	942	6	AX241501	AX241501 Sequence
c 28	242.2	49.7	166267	2	AC025942	AC025942 Homo sapi
29	238	48.9	709	10	AY074075	AY074075 Mus muscu
30	237.4	48.7	930	10	AY073772	AY073772 Mus muscu
31	236	48.5	958	6	AX241532	AX241532 Sequence
32	236	48.5	166267	2	AC025942	AC025942 Homo sapi
33	229.4	47.1	972	6	AX241530	AX241530 Sequence
34	227.8	46.8	927	10	AY073771	AY073771 Mus muscu
35	225.2	46.2	930	10	AY073613	AY073613 Mus muscu
36	224.6	46.1	927	6	AX241526	AX241526 Sequence
37	220.4	45.3	816	10	AY074144	AY074144 Mus muscu
38	217	44.6	940	6	AX241523	AX241523 Sequence
39	210.2	43.2	944	6	AX241521	AX241521 Sequence
40	206.8	42.5	483	10	AY074077	AY074077 Mus muscu
41	202.2	41.5	909	10	AY073698	AY073698 Mus muscu
42	201	41.3	931	6	AX241503	AX241503 Sequence
43	189.4	38.9	927	10	AY073090	AY073090 Mus muscu
44	187.8	38.6	918	10	AY074076	AY074076 Mus muscu
45	171.8	35.3	927	10	AY073086	AY073086 Mus muscu

ALIGNMENTS

RESULT 1
AX181436
LOCUS AX181436 487 bp DNA linear PAT 07-AUG-2001
DEFINITION Sequence 224 from Patent WO0146262.
ACCESSION AX181436
VERSION AX181436.1 GI:15141548
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 487)
AUTHORS Rouquier,S. and Giorgi,D.
TITLE Olfactory receptor genes and pseudogenes in primates and mouse
JOURNAL Patent: WO 0146262-A 224 28-JUN-2001;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES Location/Qualifiers
source 1. .487
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Taxon = 9606; gene = HSA1; Accession DDBJ/EMBL/GenBank = AF179759"
CDS <2. .>487

```
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="CAC50218.1"
                     /db_xref="GI:15141549"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDI
                     LKKKSEKGRSKAFSTCGAHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIP
                     LL"
BASE COUNT      131 a    105 c     77 g    174 t
ORIGIN


  Query Match           100.0%;  Score 487;  DB 6;  Length 487;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-98;
  Matches  487;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGTATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATAATATTTATTTTTGG 240

Qy    241 tgcttttatacaaatacccacttaatgactatcataatctcttatactcgtgtgctctt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACTATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCGGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACCAAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCC 480

Qy    481 cctgcta 487
          |||||||
Db    481 CCTGCTA 487


RESULT   2
AF179759
LOCUS       AF179759                 487 bp    DNA     linear   PRI 09-MAR-2000
DEFINITION  Homo sapiens olfactory receptor (HSA1) gene, partial cds.
ACCESSION   AF179759
VERSION     AF179759.1  GI:7211526
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Rouquier,S., Blancher,A. and Giorgi,D.
  TITLE     The olfactory receptor gene repertoire in primates and mouse:
            evidence for reduction of the functional fraction in primates
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
  MEDLINE   20183981
REFERENCE   2  (bases 1 to 487)
  AUTHORS   Giorgi,D.G. and Rouquier,S.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
            1142, rue de la Cardonille, Montpellier Cedex 5 34396, France
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            <1. .>487
                     /gene="HSA1"
     CDS             <1. .>487
                     /gene="HSA1"
                     /codon_start=2
                     /product="olfactory receptor"
                     /protein_id="AAF40348.1"
                     /db_xref="GI:7211527"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDI
                     LKKKSEKGRSKAFSTCGAHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIP
                     LL"
BASE COUNT      131 a    105 c     77 g    174 t
ORIGIN


  Query Match           100.0%;  Score 487;  DB 9;  Length 487;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-98;
  Matches  487;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGTATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATAATATTTATTTTTGG 240

Qy    241 tgcttttatacaaatacccacttaatgactatcataatctcttatactcgtgtgctctt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACTATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCGGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACCAAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCC 480

Qy    481 cctgcta 487
          |||||||
Db    481 CCTGCTA 487


RESULT   3
AX242110
LOCUS       AX242110                 486 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 858 from Patent WO0127158.
ACCESSION   AX242110
VERSION     AX242110.1  GI:15798985
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
```

```
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 858 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g708 nucleotide)"
BASE COUNT      131 a    105 c     77 g    173 t
ORIGIN


  Query Match           99.8%;  Score 486;  DB 6;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-98;
  Matches  486;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCAG 60

Qy     62 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTGCTAAGTATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACTA 120

Qy    122 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttacaa 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTACAA 180

Qy    182 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATAATATTTATTTTTGGT 240

Qy    242 gcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctcttt 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCTTTTATACAAATACCCACTTTAATGACTATCATAATCTCTTATACTCGTGTGCTCTTT 300

Qy    302 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCGGCGCC 360

Qy    362 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGCA 420

Qy    422 tctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattccc 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCTGGCTTAGCTGAAGACCAAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCCC 480

Qy    482 ctgcta 487
          ||||||
Db    481 CTGCTA 486


RESULT   4
AX181394
LOCUS       AX181394                 487 bp    DNA     linear   PAT 07-AUG-2001
DEFINITION  Sequence 182 from Patent WO0146262.
ACCESSION   AX181394
VERSION     AX181394.1  GI:15141512
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Rouquier,S. and Giorgi,D.
  TITLE     Olfactory receptor genes and pseudogenes in primates and mouse
  JOURNAL   Patent: WO 0146262-A 182 28-JUN-2001;
            CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /note="Taxon = 9598; gene = PTR210; Accession
                     DDBJ/EMBL/GenBank = AF1797"
     CDS             <2. .>487
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="CAC50200.1"
                     /db_xref="GI:15141513"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALMIFIFGAFIQIPTLMTIIISYSRVLFDI
                     LKKKSEKGRSKAFSTCSAHLLSVSLYYGTLIFMYVRPASGLAEDPDKVYSLFYTIIIP
                     LL"
BASE COUNT      129 a    107 c     78 g    173 t
ORIGIN


  Query Match           98.0%;  Score 477.4;  DB 6;  Length 487;
  Best Local Similarity   98.8%;  Pred. No. 2e-96;
  Matches  481;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy    241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATTCTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACCCAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCC 480

Qy    481 cctgcta 487
          |||||||
Db    481 CCTGCTA 487


RESULT   5
AF179735
LOCUS       AF179735                 487 bp    DNA     linear   PRI 09-MAR-2000
DEFINITION  Pan troglodytes olfactory receptor (PTR210) gene, partial cds.
ACCESSION   AF179735
VERSION     AF179735.1  GI:7211484
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Rouquier,S., Blancher,A. and Giorgi,D.
  TITLE     The olfactory receptor gene repertoire in primates and mouse:
```

```
          evidence for reduction of the functional fraction in primates
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
  MEDLINE   20183981
REFERENCE   2  (bases 1 to 487)
  AUTHORS   Giorgi,D.G. and Rouquier,S.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
            1142, rue de la Cardonille, Montpellier Cedex 5 34396, France
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     gene            <1. .>487
                     /gene="PTR210"
     CDS             <1. .>487
                     /gene="PTR210"
                     /codon_start=2
                     /product="olfactory receptor"
                     /protein_id="AAF40330.1"
                     /db_xref="GI:7211485"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALMIFIFGAFIQIPTLMTIIISYSRVLFDI
                     LKKKSEKGRSKAFSTCSAHLLSVSLYYGTLIFMYVRPASGLAEDPDKVYSLFYTIIIP
                     LL"
BASE COUNT      129 a    107 c     78 g    173 t
ORIGIN


  Query Match             98.0%;  Score 477.4;  DB 9;  Length 487;
  Best Local Similarity   98.8%;  Pred. No. 2e-96;
  Matches  481;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy    241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATTCTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACCCAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCC 480

Qy    481 cctgcta 487
          |||||||
Db    481 CCTGCTA 485


RESULT   6
AX181383
LOCUS       AX181383                485 bp    DNA     linear   PAT 07-AUG-2001
DEFINITION  Sequence 171 from Patent WO0146262.
ACCESSION   AX181383
VERSION     AX181383.1  GI:15132985
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Rouquier,S. and Giorgi,D.
  TITLE     Olfactory receptor genes and pseudogenes in primates and mouse
  JOURNAL   Patent: WO 0146262-A 171 28-JUN-2001;
            CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     misc_feature    1. .485
                     /note="Taxon = 9598; gene = PTR204; pseudogene; Accession
                     DDBJ/EMBL/GenB a"
BASE COUNT      130 a    107 c     77 g    171 t
ORIGIN


  Query Match             95.5%;  Score 465;  DB 6;  Length 485;
  Best Local Similarity   98.6%;  Pred. No. 1.2e-93;
  Matches  480;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    1;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy    241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||||||||||||||||| ||||||||||||||||  ||||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACCCAGACAAAGTGTATTCTC--TTTTACACGATTATAATTCC 478

Qy    481 cctgcta 487
          |||||||
Db    479 CCTGCTA 485


RESULT   7
AF179729
LOCUS       AF179729                485 bp    DNA     linear   PRI 09-MAR-2000
DEFINITION  Pan troglodytes PTR204 pseudogene, partial sequence.
ACCESSION   AF179729
VERSION     AF179729.1  GI:7211473
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1  (bases 1 to 485)
  AUTHORS  Rouquier,S., Blancher,A. and Giorgi,D.
  TITLE    The olfactory receptor gene repertoire in primates and mouse:
           evidence for reduction of the functional fraction in primates
  JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
  MEDLINE  20183981
REFERENCE  2  (bases 1 to 485)
  AUTHORS  Giorgi,D.G. and Rouquier,S.P.
  TITLE    Direct Submission
  JOURNAL  Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
           1142, rue de la Cardonille, Montpellier Cedex 5 34396, France
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     gene            <1. .>485
                     /gene="PTR204"
                     /pseudo
BASE COUNT      130 a    107 c     77 g    171 t
ORIGIN
```

Query Match 95.5%; Score 465; DB 9; Length 485;
Best Local Similarity 98.6%; Pred. No. 1.2e-93;
Matches 480; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db       1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db      61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttaca 180
Db     121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db     181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db     241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db     301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db     361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db     421 ATCTGGCTTAGCTGAAGACCCAGACAAAGTGTATTCTC--TTTTACACGATTATAATTCC 478

Qy     481 cctgcta 487
Db     479 CCTGCTA 485
```

```
RESULT   8
AX241993
LOCUS       AX241993                988 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 741 from Patent WO0127158.
ACCESSION   AX241993
VERSION     AX241993.1  GI:15798868
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 988)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 741 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .988
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g590 nucleotide)"
BASE COUNT      250 a    231 c    165 g    342 t
ORIGIN
```

Query Match 65.5%; Score 319; DB 6; Length 988;
Best Local Similarity 78.4%; Pred. No. 3.3e-61;
Matches 382; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```
Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db     370 TGTAGCCATATGCAATCCCTTGCTTTATCCAGTGGTGATGTCCAATAGCCTCTGTACTCA 429

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db     430 GTTTATAGGTATTTCATATTTTATTGGTTTTCTGCATTCAGCGATTCATGTGGGTTTGTT 489

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttaca 180
Db     490 ATTTAGATTAACTTTCTGCAGGTCCAATATTATACATTATTTCTACTGTGAAATTTTACA 549

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db     550 GCTGTTCAAAATTTCTTGCACCAATCCTACAGTTAATATACTTCTGATTTTCATCTTTTC 609

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db     610 AGCATTTATACAAGTCTTCACTTTTATGACTCTTATCGTCTCTTACTCCTATATTCTCTC 669

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db     670 TGCCATCCTGAAAAAGAAGTCTGAGAAGGGTAGAAGCAAAGCCTTCTCTACTTGCAGTGC 729

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db     730 CCATCTGCTCTCTGTCTCTTTGTTCTACGGCACCCTCTTCTTCATGTATGTGAGTTCTAG 789

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db     790 GTCTGGATCAGCTGCAGATCAGGCCAAAATGTATTCTTTATTTTACACAATAATAATTCC 849

Qy     481 cctgcta 487
Db     850 TTTACTA 856
```

```
RESULT   9
AC024892
LOCUS       AC024892              150083 bp    DNA     linear   HTG 15-OCT-2001
DEFINITION  Homo sapiens chromosome 3 clone RP11-214N20, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC024892
VERSION     AC024892.11  GI:16117893
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 150083)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
```

Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 150083)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (02-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Oct 14, 2001 this sequence version replaced gi:9438337.

```
-------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: HAIU
    Center clone name: RP11-214N20
-------------- Summary Statistics
    Sequencing vector: M13; L08821
    Chemistry: Dye-primer Bodipy: 2% of reads
    Chemistry: Dye-terminator Big Dye: 98% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 137907 bases at least Q40
    Consensus quality: 146103 bases at least Q30
    Consensus quality: 149133 bases at least Q20
    Estimated insert size: 151155; sum-of-contigs estimation
    Quality coverage: 0x in Q20 bases; agarose-fp estimation
    Quality coverage: 4.1x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    34152: contig of 34152 bp in length
*    34153    34252: gap of unknown length
*    34253    55507: contig of 21255 bp in length
*    55508    55607: gap of unknown length
*    55608    79422: contig of 23815 bp in length
*    79423    79522: gap of unknown length
*    79523    94794: contig of 15272 bp in length
*    94795    94894: gap of unknown length
*    94895   107021: contig of 12127 bp in length
*   107022   107121: gap of unknown length
*   107122   116733: contig of 9612 bp in length
*   116734   116833: gap of unknown length
*   116834   125530: contig of 8697 bp in length
*   125531   125630: gap of unknown length
*   125631   135300: contig of 9670 bp in length
*   135301   135400: gap of unknown length
*   135401   140864: contig of 5464 bp in length
*   140865   140964: gap of unknown length
*   140965   146546: contig of 5582 bp in length
*   146547   146646: gap of unknown length
*   146647   150083: contig of 3437 bp in length.
```

FEATURES Location/Qualifiers

source 1. .150083
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-214N20"

BASE COUNT 45301 a 29114 c 27778 g 46869 t 1021 others

ORIGIN

Query Match 65.5%; Score 319; DB 2; Length 150083;
Best Local Similarity 78.4%; Pred. No. 1.9e-61;
Matches 382; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          |||||||||||| |||||||||||||||||||| |||||||||| |  ||| |  ||||
Db  57586 TGTAGCCATATGCAATCCCTTGCTTTATCCAGTGGTGATGTCCAATAGCCTCTGTACTCA 57645

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          |||  || |||||||||||| | |||||||| |||||| |   | |||||||| ||||  |
Db  57646 GTTTATAGGTATTTCATATTTTATTGGTTTTCTGCATTCAGCGATTCATGTGGGTTTGTT 57705

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
          |||  || ||||||||||||||||  || || ||||||||||||||||||||||||||||||
Db  57706 ATTTAGATTAACTTTCTGCAGGTCCAATATTATACATTATTTCTACTGTGAAATTTTACA 57765

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||| ||||    |||  |  ||||   |||  | || || || |||
Db  57766 GCTGTTCAAAATTTCTTGCACCAATCCTACAGTTAATATACTTCTGATTTTCATCTTTTC 57825

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          || ||||||||| |   |||||| |||||| | ||  |||||||  |   | | ||||
Db  57826 AGCATTTATACAAGTCTTCACTTTTATGACTCTTATCGTCTCTTACTCCTATATTCTCTC 57885

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          ||  || |||||||| |||||||| ||||| |||||||||||||||| || ||| | ||
Db  57886 TGCCATCCTGAAAAAGAAGTCTGAGAAGGGTAGAAGCAAAGCCTTCTCTACTTGCAGTGC 57945

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          |||||||||| |||||||| |||| |||||| || ||  |||||||||||||| || ||
Db  57946 CCATCTGCTCTCTGTCTCTTTGTTCTACGGCACCCTCTTCTTCATGTATGTGAGTTCTAG 58005

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||||| | ||||| ||| || | |||| |||||||| | |||||||| || ||||||||
Db  58006 GTCTGGATCAGCTGCAGATCAGGCCAAAATGTATTCTTTATTTTACACAATAATAATTCC 58065

Qy    481 cctgcta 487
            | |||
Db  58066 TTTACTA 58072
```

```
RESULT  10
AC105564/c
LOCUS       AC105564              138556 bp    DNA     linear   HTG 09-JAN-2002
DEFINITION  Rattus norvegicus clone CH230-250G11, *** SEQUENCING IN PROGRESS
            ***, 60 unordered pieces.
ACCESSION   AC105564
VERSION     AC105564.1  GI:18092786
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 138556)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 138556)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
              Center project name: GNEZ
              Center clone name: CH230-250G11
            -------------- Summary Statistics
              Assembly program: Phrap; version 0.990329First call to
              findPhrapList
              Consensus quality: 104982 bases at least Q40
              Consensus quality: 113437 bases at least Q30
              Consensus quality: 120052 bases at least Q20
              Estimated insert size: 98776; sum-of-contigs estimation
              Quality coverage: 0x in Q20 bases; agarose-fp estimation
              Quality coverage: 1.3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 60 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    12026: contig of 12026 bp in length
            *    12027    12126: gap of unknown length
            *    12127    17583: contig of 5457 bp in length
            *    17584    17683: gap of unknown length
            *    17684    22927: contig of 5244 bp in length
            *    22928    23027: gap of unknown length
            *    23028    27987: contig of 4960 bp in length
            *    27988    28087: gap of unknown length
            *    28088    32504: contig of 4417 bp in length
            *    32505    32604: gap of unknown length
            *    32605    35333: contig of 2729 bp in length
            *    35334    35433: gap of unknown length
            *    35434    37327: contig of 1894 bp in length
            *    37328    37427: gap of unknown length
            *    37428    40912: contig of 3485 bp in length
            *    40913    41012: gap of unknown length
            *    41013    44129: contig of 3117 bp in length
            *    44130    44229: gap of unknown length
            *    44230    46597: contig of 2368 bp in length
            *    46598    46697: gap of unknown length
            *    46698    50418: contig of 3721 bp in length
            *    50419    50518: gap of unknown length
            *    50519    52990: contig of 2472 bp in length
            *    52991    53090: gap of unknown length
            *    53091    56522: contig of 3432 bp in length
            *    56523    56622: gap of unknown length
            *    56623    59005: contig of 2383 bp in length
            *    59006    59105: gap of unknown length
            *    59106    61667: contig of 2562 bp in length
            *    61668    61767: gap of unknown length
            *    61768    63899: contig of 2132 bp in length
            *    63900    63999: gap of unknown length
            *    64000    68087: contig of 4088 bp in length
            *    68088    68187: gap of unknown length
            *    68188    70694: contig of 2507 bp in length
            *    70695    70794: gap of unknown length
            *    70795    73386: contig of 2592 bp in length
            *    73387    73486: gap of unknown length
            *    73487    75692: contig of 2206 bp in length
            *    75693    75792: gap of unknown length
            *    75793    77929: contig of 2137 bp in length
            *    77930    78029: gap of unknown length
            *    78030    79296: contig of 1267 bp in length
            *    79297    79396: gap of unknown length
            *    79397    82238: contig of 2842 bp in length
            *    82239    82338: gap of unknown length
            *    82339    84658: contig of 2320 bp in length
            *    84659    84758: gap of unknown length
            *    84759    87336: contig of 2578 bp in length
            *    87337    87436: gap of unknown length
            *    87437    89496: contig of 2060 bp in length
            *    89497    89596: gap of unknown length
            *    89597    90917: contig of 1321 bp in length
            *    90918    91017: gap of unknown length
            *    91018    92625: contig of 1608 bp in length
```

```
          *    92626    92725: gap of unknown length
          *    92726    93955: contig of 1230 bp in length
          *    93956    94055: gap of unknown length
          *    94056    95238: contig of 1183 bp in length
          *    95239    95338: gap of unknown length
          *    95339    96506: contig of 1168 bp in length
          *    96507    96606: gap of unknown length
          *    96607    98414: contig of 1808 bp in length
          *    98415    98514: gap of unknown length
          *    98515   100270: contig of 1756 bp in length
          *   100271   100370: gap of unknown length
          *   100371   101914: contig of 1544 bp in length
          *   101915   102014: gap of unknown length
          *   102015   103378: contig of 1364 bp in length
          *   103379   103478: gap of unknown length
          *   103479   104660: contig of 1182 bp in length
          *   104661   104760: gap of unknown length
          *   104761   106384: contig of 1624 bp in length
          *   106385   106484: gap of unknown length
          *   106485   107638: contig of 1154 bp in length
          *   107639   107738: gap of unknown length
          *   107739   109065: contig of 1327 bp in length
          *   109066   109165: gap of unknown length
          *   109166   110732: contig of 1567 bp in length
          *   110733   110832: gap of unknown length
          *   110833   112101: contig of 1269 bp in length
          *   112102   112201: gap of unknown length
          *   112202   113509: contig of 1308 bp in length
          *   113510   113609: gap of unknown length
          *   113610   115100: contig of 1491 bp in length
          *   115101   115200: gap of unknown length
          *   115201   116535: contig of 1335 bp in length
          *   116536   116635: gap of unknown length
          *   116636   118099: contig of 1464 bp in length
          *   118100   118199: gap of unknown length
          *   118200   119225: contig of 1026 bp in length
          *   119226   119325: gap of unknown length
          *   119326   121099: contig of 1774 bp in length
          *   121100   121199: gap of unknown length
          *   121200   122556: contig of 1357 bp in length
          *   122557   122656: gap of unknown length
          *   122657   124309: contig of 1653 bp in length
          *   124310   124409: gap of unknown length
          *   124410   125647: contig of 1238 bp in length
          *   125648   125747: gap of unknown length
          *   125748   126786: contig of 1039 bp in length
          *   126787   126886: gap of unknown length
          *   126887   128160: contig of 1274 bp in length
          *   128161   128260: gap of unknown length
          *   128261   129777: contig of 1517 bp in length
          *   129778   129877: gap of unknown length
          *   129878   130976: contig of 1099 bp in length
          *   130977   131076: gap of unknown length
          *   131077   132445: contig of 1369 bp in length
          *   132446   132545: gap of unknown length
          *   132546   133844: contig of 1299 bp in length
          *   133845   133944: gap of unknown length
          *   133945   135012: contig of 1068 bp in length

  Query Match             60.4%;  Score 294;  DB 2;  Length 138556;
  Best Local Similarity   75.2%;  Pred. No. 6.9e-56;
  Matches  366;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||||||||||||||| ||  |||| |||| |||| |||||||||| | |||| || ||
Db  131755 TGTAGCCATATGTAACCCTCTGCTCTATCTAGTGGTGATGTCCAAAAGACTCTGCACTGT 131696

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           | |  | ||||| |||||| | |||||||| ||  | ||| | |||||||||| | ||| |
Db  131695 GCTAATCAGTATGTCATATATCATTGGTTTTCTAAACCCTATAGTTCATGTGGGATTATT 131636

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           |||  || |||| || || |||| |||| | ||| || ||||||||||||||||||| |

Db  131635 ATTCAGATTAACATTTTGTAGGTCTAACGTTATAGATCATTTCTACTGTGAAATTTTGCC 131576

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||| |  | |||||| |||| | |||| || |||| ||| |  ||||  |||||||||
Db  131575 ACTCTATACAATTTCCTGCACAGATCCAACTCTTAATGCATTGGTAATTNTTATTTTTGC 131516

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           | |||  ||||||||    || |   |||| ||  || |||||||| |||| || || ||
Db  131515 TTCTTCCATACAAATCAGTACCTCTGTGACCATTGTAGTCTCTTATGCTCGAGTCCTATT 131456

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           || | ||||| | |  |||||||| |||||||||| ||||| |||| ||| ||| |  |
Db  131455 TGCTGTTCTGTACATGAAGTCTGAGAAGGGCAGAAGAAAAGCTTTCTTCACCTGCAGTTC 131396

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           |||  || | |||||||| |||| ||| || |||||  ||||||||||||  ||| ||
Db  131395 CCACTTGTTCTCTGTCTCTTTGTTCTATGGTACTCTCCTCTTCATGTATGTAAGTCATGG 131336

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           |||| |  || |||| |||| || ||| ||||||||||||| || ||  || | |||||
Db  131335 GTCTGCCCCAGGTGAAAACCAGGATAAAATGTATTCTCTGTTCTATACAGTTGTGATTCC 131276

Qy     481 cctgcta 487
           |||||||
Db  131275 TCTGCTA 131269


RESULT  11
AY074239
LOCUS       AY074239                693 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR182-11P pseudogene, partial
            sequence.
ACCESSION   AY074239
VERSION     AY074239.1  GI:18481401
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
   PUBMED   11802173
REFERENCE   2  (bases 1 to 693)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            <1. .>693
                     /gene="olfactory receptor MOR182-11P"
                     /note="potential coding region disrupted by one or more
                     internal stop codons"
                     /pseudo
                     /evidence=not_experimental
BASE COUNT      177 a    139 c    125 g    252 t
ORIGIN


  Query Match             59.9%;  Score 291.8;  DB 10;  Length 693;
  Best Local Similarity   74.9%;  Pred. No. 3.8e-55;
  Matches  365;  Conservative    0;  Mismatches  122;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           |||||||||||| || ||  |||| |||| |||| |||||||||| | |||| || ||
Db     144 TGTAGCCATATGCAACCCTCTGCTCTATCTAGTGGTGATGTCCAAGAGACTCTGCACTGT 203

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           |||  | ||| | |||||| |||||||| | ||  ||||| | |||||||| | ||| |
```

```
Db    204 GTTCATCAGTTTATCATATATAATTGGTGTTCTAAATCCTATAGTTCATGTGGGATTATT 263

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
          |||  || |||| ||||| |  | |||  | ||  || |||||||||||||||| || |
Db    264 ATTCAGATTAACATTCTGTAAATCTAATGTTATTGATCATTTCTACTGTGAAATCTTGCC 323

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          ||| | || |||||| ||||  | | ||||| | || ||| |  |||| ||||||||||
Db    324 ACTCTACACAATTTCTTGCACAGATGCATCTCTGAATGCATTGGTAATTTTTATTTTTGC 383

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          | |||  |||||||     || |   ||||||| ||| |||||||| | ||||| || ||
Db    384 TTCTTCCATACAAATCAGTACCTCTGTGACTATTATAGTCTCTTATGCCCGTGTCCTATT 443

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |    | ||||| |  |||||||| |||||||||| ||||| ||||  || ||| | ||
Db    444 TTCGGTCCTGAACATGAAGTCTGAGAAGGGCAGAAGAAAAGCTTTCTTTACCTGCAGTGC 503

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||| ||||| |||||||| |||| ||| || || ||  |||||||||||||| ||| ||
Db    504 CCACCTGCTCTCTGTCTCTTTGTTCTATGGTACCCTCCTCTTCATGTATGTGAGTCATGG 563

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
            ||||    || |||| |||| || ||| ||||||||||||| |||||  || | |||||
Db    564 GTCTGCACCAGGTGAAAACCAGGATAAAATGTATTCTCTGTTCTACACAGTTGTGATTCC 623

Qy    481 cctgcta 487
            |||||||
Db    624 TCTGCTA 630


RESULT  12
AF091575
LOCUS       AF091575                663 bp    mRNA    linear   ROD 20-OCT-1998
DEFINITION  Rattus norvegicus isolate HFV-FD1 olfactory receptor mRNA, partial
            cds.
ACCESSION   AF091575
VERSION     AF091575.1  GI:3769640
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 663)
  AUTHORS   Singer,M.S., Hughes,T.E., Shepherd,G.M. and Greer,C.A.
  TITLE     Isolation of olfactory receptor mRNA sequences from olfactory bulb
            glomerular layer
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 663)
  AUTHORS   Singer,M.S., Hughes,T.E., Shepherd,G.M. and Greer,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-1998) Section of Neurobiology, Yale University
            School of Medicine, 236 FMB, 333 Cedar Street, New Haven, CT 06510,
            USA
FEATURES             Location/Qualifiers
     source          1. .663
                     /organism="Rattus norvegicus"
                     /isolate="HFV-FD1"
                     /db_xref="taxon:10116"
     CDS             <1. .>663
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="AAC64595.1"
                     /db_xref="GI:3769641"
                     /translation="GNLAFADACTSSSVTPKMLMKFLNRNDMISMGECFAQYYFFCFS
                     ATTEIFLLVAMAYDRYVAICNPLLYLVVMSKRLCTVLISMSYIIGFLNPIVHVGLLFR
                     LTFCRSNVIDHFYCEILPLYTISCTDPTLNALVIFIFASSIQISTSVTIVVSYARVLF
                     AVLYMKSEKGRRKAFFTCSSHLFSVSLFYGTLLFMYVSHGSAPGENQDKMYSLFYTVV
                     IPL"
BASE COUNT      161 a    137 c    119 g    246 t
ORIGIN
```

```
  Query Match           59.6%;  Score 290.4;  DB 10;  Length 663;
  Best Local Similarity   75.0%;  Pred. No. 7.9e-55;
  Matches  363;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          |||||||||||||||| ||  |||| |||| |||| |||||||||| | |||| || ||
Db    180 TGTAGCCATATGTAACCCTCTGCTCTATCTAGTGGTGATGTCCAAAAGACTCTGCACTGT 239

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | |  | ||||| |||||| | |||||||| ||  | ||| | ||||||||| | ||| |
Db    240 GCTAATCAGTATGTCATATATCATTGGTTTTCTAAACCCTATAGTTCATGTGGGATTATT 299

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
          |||  || |||| || || |||| |||| | ||  || ||||||||||||||||||| |
Db    300 ATTCAGATTAACATTTTGTAGGTCTAACGTTATTGATCATTTCTACTGTGAAATTTTGCC 359

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          ||| |  | |||||| ||||  | |||| || |||| ||| |  |||| ||||||||||
Db    360 ACTCTATACAATTTCCTGCACAGATCCAACTCTTAATGCATTGGTAATTTTTATTTTTGC 419

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          | |||  ||||||||     || |   |||| ||  || |||||||| |||| || || ||
Db    420 TTCTTCCATACAAATCAGTACCTCTGTGACCATTGTAGTCTCTTATGCTCGAGTCCTATT 479

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          || | ||||| | |  |||||||| |||||||||| ||||| |||| ||| ||| |  |
Db    480 TGCTGTTCTGTACATGAAGTCTGAGAAGGGCAGAAGAAAAGCTTTCTTCACCTGCAGTTC 539

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          |||  || | |||||||| |||| ||| || |||||  |||||||||||||  ||| ||
Db    540 CCACTTGTTCTCTGTCTCTTTGTTCTATGGTACTCTCCTCTTCATGTATGTAAGTCATGG 599

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||| |  || |||| |||| || ||| ||||||||||||| || ||  || | |||||
Db    600 GTCTGCCCCAGGTGAAAACCAGGATAAAATGTATTCTCTGTTCTATACAGTTGTGATTCC 659

Qy    481 cctg 484
          ||||
Db    660 TCTG 663


RESULT  13
AY073096
LOCUS       AY073096                918 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR182-3 gene, complete cds.
ACCESSION   AY073096
VERSION     AY073096.1  GI:18479489
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
   PUBMED   11802173
REFERENCE   2  (bases 1 to 918)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     mRNA            <1. .>918
                     /product="olfactory receptor MOR182-3"
     CDS             1. .918
                     /codon_start=1
                     /evidence=not_experimental
```

```
                     /product="olfactory receptor MOR182-3"
                     /protein_id="AAL60759.1"
                     /db_xref="GI:18479490"
                     /translation="MEKNSTLMTEFVLRGITDRPELQVPLFLVFFFIYVTTMVGNLGL
                     IFLIWKDPHLHTPMYLFLGNLAFADACTSSSVTPKMLMKFLNKNDIISRGECFAQFYF
                     FCFSATTEIFLLVAMAYDRYVAICNPLLYLVVMSKRLCTVLISLSYIIGVLNPIVHVG
                     LLFRLTFCKSNVIDHFYCEIMPLYAISCTDPSLNGLVAFIFASSIQISTSVTIVVSYA
                     RVLFAVLNMKSERGRRKAFFTCSAHLLSVSLFYGTLLFMYVSPGSGSGKQKDKMYSLF
                     YTVVIPLLNPFIYSLRNKEVLGALKKLIK"
BASE COUNT      228 a    202 c    170 g    318 t
ORIGIN


  Query Match             57.9%;  Score 282.2;  DB 10;  Length 918;
  Best Local Similarity   73.7%;  Pred. No. 5.1e-53;
  Matches  359;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           |||||||||||| || ||  |||| |||| |||| |||||||||| | |||| |  ||
Db     366 TGTAGCCATATGCAACCCTCTGCTCTATCTAGTGGTGATGTCCAAGAGACTCTGTACTGT 425

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           |||  | ||| | |||||| |||||||| | ||  ||||| | |||||||| | ||| |
Db     426 GTTAATCAGTTTATCATATATAATTGGTGTTCTAAATCCTATAGTTCATGTGGGATTATT 485

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           |||  || |||| ||||| |  | |||  | ||  || ||||||||||||||||  | |
Db     486 ATTCAGATTAACATTCTGTAAATCTAATGTTATTGATCATTTCTACTGTGAAATCATGCC 545

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||| |     ||||| |||  | ||||||| |||| | | |   |    ||||||||||
Db     546 ACTCTATGCGATTTCTTGCACAGATCCATCTCTTAATGGATTGGTGGCTTTTATTTTTGC 605

Qy     241 tgcttttatacaaatacccacttaatgactatcataatctcttatactcgtgtgctctt 300
           | ||| |||||||||    || |   |||||||| || |||||||| | ||||| || ||
Db     606 TTCTTCTATACAAATCAGTACCTCTGTGACTATCGTAGTCTCTTATGCGCGTGTCCTTTT 665

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           || | | ||||| |  |||||||| | ||||||||| ||||| || | ||| ||| | ||
Db     666 TGCTGTCCTGAACATGAAGTCTGAGAGGGGCAGAAGAAAAGCTTTTTTCACCTGCAGTGC 725

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||| ||||| |||||||| |||| ||| || || ||  |||||||||||||| | ||||
Db     726 CCACCTGCTCTCTGTCTCTTTGTTCTATGGTACCCTCCTCTTCATGTATGTGAGCCCTGG 785

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           |||||  | || | || |  | || ||| |||||||||||| |||||  || | |||||
Db     786 GTCTGGATCAGGTAAACAAAAGGATAAAATGTATTCTCTGTTCTACACAGTTGTGATTCC 845

Qy     481 cctgcta 487
           |||||||
Db     846 TCTGCTA 852


RESULT  14
AY073611
LOCUS       AY073611                 921 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR182-5 gene, complete cds.
ACCESSION   AY073611
VERSION     AY073611.1  GI:18480519
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 921)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
   PUBMED   11802173
REFERENCE   2  (bases 1 to 921)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .921
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     mRNA            <1. .>921
                     /product="olfactory receptor MOR182-5"
     CDS             1. .921
                     /codon_start=1
                     /evidence=not_experimental
                     /product="olfactory receptor MOR182-5"
                     /protein_id="AAL61274.1"
                     /db_xref="GI:18480520"
                     /translation="MEVNRTLVTAFILRGITDLPELQVPMFLVFFFIYVTTMVGNLGL
                     IVLIWKDPRLHTPMYFFLGSLAFADACTSSSVTPRMLVNILDNGKMISLSECMAQYYV
                     FGSSATTECFLLVAMAYDRYVAICNPLLYLVVMSNRVCTCLISGSYIIGFLHPLIHVG
                     LLFRLTFCKSNIIDHFYCEILPLYTISCTDPSINAFVVFIFAAVIQAVTFMSIAVSYA
                     HVLFSILKTKSEKGRRKAFSTCSAHLLSVSLFYGTLFFMYVSPGSGPSKYKNKMYSLF
                     YTIVIPLLNPFIYSLRNKEVLGALRKIMKP"
BASE COUNT      224 a    218 c    173 g    306 t
ORIGIN


  Query Match             57.9%;  Score 282.2;  DB 10;  Length 921;
  Best Local Similarity   73.7%;  Pred. No. 5.1e-53;
  Matches  359;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           |||||||||||| ||||| ||||  ||| ||||||||||| | |  || ||
Db     366 TGTAGCCATATGCAATCCTTTGCTCTATCTTGTGGTGATGTCCAACAGAGTGTGCACTTG 425

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
             || |||||   |||||| |||| ||||| |||||||| ||  |||||||  || ||||
Db     426 CCTGATAAGTGGCTCATATATAATCGGTTTTCTGCATCCACTTATTCATGTAGGTCTACT 485

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           |||  || |||| |||||||  |  || |||||| || |||| ||||||||  ||  | |
Db     486 ATTTAGATTAACCTTCTGCAAATCCAATATAATAGATCATTTTTACTGTGAGATCCTGCC 545

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||| |  | |||||| ||||  |  |||||||||||| ||| |  |  | || ||||||
Db     546 ACTCTATACAATTTCTTGCACCGACCCATCTATTAATGCATTTGTGGTCTTCATTTTTGC 605

Qy     241 tgcttttatacaaatacccacttaatgactatcataatctcttatactcgtgtgctctt 300
           |||| | ||||||       || || |||| |||   | |||| ||| | |  || |||||
Db     606 TGCTGTGATACAAGCTGTTACCTTCATGAGTATTGCAGTCTCCTATGCCCACGTCCTCTT 665

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           |   || ||||||| ||||||||| |||||||||||| |||||||||||||| ||| | ||
Db     666 TTCCATCCTGAAAACAAAGTCTGAGAAGGGCAGAAGAAAAGCCTTCTCCACCTGCAGTGC 725

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||| || || |||||||| ||| ||| |||||||| |||| |||| ||| | ||||
Db     726 CCACCTACTCTCTGTCTCTCTGTTCTATGGAACTCTCTTCTTTATGTACGTGAGCCCTGG 785

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           |||||   |  | || |  |  | ||  ||||||||||||||||||| ||| | |||||
Db     786 GTCTGGACCAAGTAAATATAAGAATAAGATGTATTCTCTGTTTTACACCATTGTGATTCC 845

Qy     481 cctgcta 487
           || |||
Db     846 TCTACTA 852


RESULT  15
AC105564
LOCUS       AC105564               138556 bp    DNA     linear   HTG 09-JAN-2002
DEFINITION  Rattus norvegicus clone CH230-250G11, *** SEQUENCING IN PROGRESS
            ***, 60 unordered pieces.
ACCESSION   AC105564
VERSION     AC105564.1  GI:18092786
```

```
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 138556)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 138556)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
              Center project name: GNEZ
              Center clone name: CH230-250G11
            -------------- Summary Statistics
              Assembly program: Phrap; version 0.990329First call to
            findPhrapList
              Consensus quality: 104982 bases at least Q40
              Consensus quality: 113437 bases at least Q30
              Consensus quality: 120052 bases at least Q20
              Estimated insert size: 98776; sum-of-contigs estimation
              Quality coverage: 0x in Q20 bases; agarose-fp estimation
              Quality coverage: 1.3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 60 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    12026: contig of 12026 bp in length
            *    12027    12126: gap of unknown length
            *    12127    17583: contig of 5457 bp in length
            *    17584    17683: gap of unknown length
            *    17684    22927: contig of 5244 bp in length
            *    22928    23027: gap of unknown length
            *    23028    27987: contig of 4960 bp in length
            *    27988    28087: gap of unknown length
            *    28088    32504: contig of 4417 bp in length
            *    32505    32604: gap of unknown length
            *    32605    35333: contig of 2729 bp in length
            *    35334    35433: gap of unknown length
            *    35434    37327: contig of 1894 bp in length
            *    37328    37427: gap of unknown length
            *    37428    40912: contig of 3485 bp in length
            *    40913    41012: gap of unknown length
            *    41013    44129: contig of 3117 bp in length
            *    44130    44229: gap of unknown length
            *    44230    46597: contig of 2368 bp in length
            *    46598    46697: gap of unknown length
            *    46698    50418: contig of 3721 bp in length
            *    50419    50518: gap of unknown length
            *    50519    52990: contig of 2472 bp in length
            *    52991    53090: gap of unknown length
            *    53091    56522: contig of 3432 bp in length
            *    56523    56622: gap of unknown length
            *    56623    59005: contig of 2383 bp in length
            *    59006    59105: gap of unknown length
            *    59106    61667: contig of 2562 bp in length
            *    61668    61767: gap of unknown length
            *    61768    63899: contig of 2132 bp in length
            *    63900    63999: gap of unknown length
            *    64000    68087: contig of 4088 bp in length
            *    68088    68187: gap of unknown length
            *    68188    70694: contig of 2507 bp in length
            *    70695    70794: gap of unknown length
            *    70795    73386: contig of 2592 bp in length
            *    73387    73486: gap of unknown length
            *    73487    75692: contig of 2206 bp in length
            *    75693    75792: gap of unknown length
            *    75793    77929: contig of 2137 bp in length
            *    77930    78029: gap of unknown length
            *    78030    79296: contig of 1267 bp in length
            *    79297    79396: gap of unknown length
            *    79397    82238: contig of 2842 bp in length
            *    82239    82338: gap of unknown length
            *    82339    84658: contig of 2320 bp in length
            *    84659    84758: gap of unknown length
            *    84759    87336: contig of 2578 bp in length
            *    87337    87436: gap of unknown length
            *    87437    89496: contig of 2060 bp in length
            *    89497    89596: gap of unknown length
            *    89597    90917: contig of 1321 bp in length
            *    90918    91017: gap of unknown length
            *    91018    92625: contig of 1608 bp in length
            *    92626    92725: gap of unknown length
            *    92726    93955: contig of 1230 bp in length
            *    93956    94055: gap of unknown length
            *    94056    95238: contig of 1183 bp in length
            *    95239    95338: gap of unknown length
            *    95339    96506: contig of 1168 bp in length
            *    96507    96606: gap of unknown length
            *    96607    98414: contig of 1808 bp in length
            *    98415    98514: gap of unknown length
```

```
  *    98515   100270: contig of 1756 bp in length
  *   100271   100370: gap of unknown length
  *   100371   101914: contig of 1544 bp in length
  *   101915   102014: gap of unknown length
  *   102015   103378: contig of 1364 bp in length
  *   103379   103478: gap of unknown length
  *   103479   104660: contig of 1182 bp in length
  *   104661   104760: gap of unknown length
  *   104761   106384: contig of 1624 bp in length
  *   106385   106484: gap of unknown length
  *   106485   107638: contig of 1154 bp in length
  *   107639   107738: gap of unknown length
  *   107739   109065: contig of 1327 bp in length
  *   109066   109165: gap of unknown length
  *   109166   110732: contig of 1567 bp in length
  *   110733   110832: gap of unknown length
  *   110833   112101: contig of 1269 bp in length
  *   112102   112201: gap of unknown length
  *   112202   113509: contig of 1308 bp in length
  *   113510   113609: gap of unknown length
  *   113610   115100: contig of 1491 bp in length
  *   115101   115200: gap of unknown length
  *   115201   116535: contig of 1335 bp in length
  *   116536   116635: gap of unknown length
  *   116636   118099: contig of 1464 bp in length
  *   118100   118199: gap of unknown length
  *   118200   119225: contig of 1026 bp in length
  *   119226   119325: gap of unknown length
  *   119326   121099: contig of 1774 bp in length
  *   121100   121199: gap of unknown length
  *   121200   122556: contig of 1357 bp in length
  *   122557   122656: gap of unknown length
  *   122657   124309: contig of 1653 bp in length
  *   124310   124409: gap of unknown length
  *   124410   125647: contig of 1238 bp in length
  *   125648   125747: gap of unknown length
  *   125748   126786: contig of 1039 bp in length
  *   126787   126886: gap of unknown length
  *   126887   128160: contig of 1274 bp in length
  *   128161   128260: gap of unknown length
  *   128261   129777: contig of 1517 bp in length
  *   129778   129877: gap of unknown length
  *   129878   130976: contig of 1099 bp in length
  *   130977   131076: gap of unknown length
  *   131077   132445: contig of 1369 bp in length
  *   132446   132545: gap of unknown length
  *   132546   133844: contig of 1299 bp in length
  *   133845   133944: gap of unknown length
  *   133945   135012: contig of 1068 bp in length

  Query Match             57.9%;  Score 282.2;  DB 2;  Length 138556;
  Best Local Similarity   73.7%;  Pred. No. 2.9e-53;
  Matches  359;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           | ||||||||||| || ||  |||| ||||  ||| ||||||||||||| | |  || ||
Db 135143 TCTAGCCATATGCAACCCTCTGCTCTATCTTGTGGTGATGTCCAACAGAGTGTGCACTTG 135202

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
             || |||||  ||||||| |||| ||||| |||||||| || ||||||||| | ||| |
Db 135203 CCTGATAAGTGGTTCATATATAATCGGTTTTCTGCATCCACTTGTTCATGTGGGATTATT 135262

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           |||  ||||||| ||||||| ||  ||  ||||| || | || |||||||| ||  | |
Db 135263 ATTTAGACTAACCTTCTGCAAGTCCAATGTAATAGATCACTTTTACTGTGAGATCCTGCC 135322

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
             || |  | |||||| |||| ||  |||||||||||| ||| |  |  | || |||||||
Db 135323 TCTCTATACAATTTCTTGCACTGACCCATCTATTAATGCATTCGTGGTCTTCATTTTTGC 135382

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           |||  | ||||||      || || |||| | |   | |||| ||| | | ||| |||||
Db 135383 TGCCGTAATACAAGCCATTACCTTCATGAGTGTTGCAGTCTCCTATGCCCATGTCCTCTT 135442

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           |   || ||||||| ||||||| | ||| | ||||| |||||||||||||| ||| | ||
Db 135443 TTCCATCCTGAAAACAAAGTCTCAGAAGAGTAGAAGAAAAGCCTTCTCCACCTGCAGTGC 135502

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||| ||||| |||||||| |||| ||| ||||||||  |||||||||||||| | ||||
Db 135503 CCACCTGCTCTCTGTCTCTTTGTTCTATGGAACTCTCTTCTTCATGTATGTGAGCCCTGG 135562

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           |||||   |  | || |  |  | ||| |||||||||||||||||||| ||| | |||||
Db 135563 GTCTGGACCAAGTAAATATAAGAATAAAATGTATTCTCTGTTTTACACCATTGTGATTCC 135622

Qy     481 cctgcta 487
           || |||
Db 135623 TCTACTA 135629


Search completed: June 27, 2002, 11:17:02
Job time: 3515 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 10:18:27 ; Search time 207.32 Seconds
(without alignments)
4033.076 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	487	100.0	487	22	AAH84016	Human olfactory re
2	486	99.8	486	22	AAH32285	Human olfactory re
3	477.4	98.0	487	22	AAH83992	Hylobates lar olfa
4	465	95.5	485	22	AAH83986	Pan troglodytes ol
5	319	65.5	988	22	AAH32168	Human olfactory re
6	275.8	56.6	953	22	AAI99556	Human expressed po
7	245.4	50.4	487	22	AAH83978	Pan troglodytes ol
8	242.2	49.7	942	22	AAH31676	Human olfactory re
9	242.2	49.7	945	22	AAS42326	Human cDNA encodin
10	236	48.5	958	22	AAH31707	Human olfactory re
11	229.4	47.1	972	22	AAH31705	Human olfactory re
12	224.6	46.1	927	22	AAH31701	Human olfactory re
13	224.6	46.1	978	22	AAS42325	Human cDNA encodin
14	217	44.6	940	22	AAH31698	Human olfactory re
15	210.2	43.2	944	22	AAH31696	Human olfactory re
16	201	41.3	931	22	AAH31678	Human olfactory re
17	165.4	34.0	942	22	AAH32121	Human olfactory re
18	161.8	33.2	2337	22	AAF30803	Human G protein co
19	160.2	32.9	1687	22	AAF30802	Human G protein co
20	158.6	32.6	936	22	AAH31786	Human olfactory re
21	158.6	32.6	945	22	AAS42245	Human cDNA encodin
22	158.6	32.6	1033	22	AAH75736	Partial NOV 2 codi
23	158.6	32.6	1040	22	AAH75717	Human NOV 2 coding
24	156.4	32.1	342	22	AAH32182	Human olfactory re
25	153.8	31.6	938	22	AAH31781	Human olfactory re
26	150	30.8	957	24	ABK16632	Human G-coupled re
27	150	30.8	984	22	AAH31820	Human olfactory re
28	150	30.8	1080	22	AAS42272	Human cDNA encodin
29	148.4	30.5	1050	22	AAD13718	Human G-protein co
30	147.8	30.3	933	22	AAH31708	Human olfactory re
31	147.8	30.3	951	22	AAS42228	Human cDNA encodin
32	147.8	30.3	951	22	AAH32010	Human olfactory re
33	147.8	30.3	954	22	AAS42357	Human cDNA encodin
34	146.8	30.1	960	22	AAH31750	Human olfactory re
35	146.2	30.0	753	22	AAH31787	Human olfactory re
36	146.2	30.0	975	22	AAS42457	Human cDNA encodin
37	145.2	29.8	144460	21	AAZ93815	Olfactory receptor
38	144.4	29.7	933	22	AAS42461	Human cDNA encodin
39	143.6	29.5	900	22	AAH32479	Human olfactory re
40	143.6	29.5	1023	22	AAS42273	Human cDNA encodin
41	140.4	28.8	756	22	AAH32142	Human olfactory re
42	140.4	28.8	948	22	AAH32140	Human olfactory re
43	140.4	28.8	975	22	AAS42390	Human cDNA encodin
44	139.4	28.6	487	22	AAH84020	Human olfactory re
45	139.4	28.6	960	22	AAH32144	Human olfactory re

ALIGNMENTS

RESULT 1
AAH84016
ID AAH84016 standard; DNA; 487 BP.
XX
AC AAH84016;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human olfactory receptor encoding gene 4.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance; ds.
XX
OS Homo sapiens.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB02017.
XX
PR 22-DEC-1999; 99US-0171746.
PR 21-DEC-2000; 2000US-0747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR P-PSDB; AAG98521.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful

PT for analysis odours e.g. in food processing and perfumery -
XX
PS Claim 1; Page 273-274; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 487 BP; 131 A; 105 C; 77 G; 174 T; 0 other;

Query Match 100.0%; Score 487; DB 22; Length 487;
Best Local Similarity 100.0%; Pred. No. 5e-116;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480

Qy    481 cctgcta 487
          |||||||
Db    481 cctgcta 487
```

RESULT 2
AAH32285
ID AAH32285 standard; DNA; 486 BP.
XX
AC AAH32285;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 858.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 523; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 486 BP; 131 A; 105 C; 77 G; 173 T; 0 other;

Query Match 99.8%; Score 486; DB 22; Length 486;
Best Local Similarity 100.0%; Pred. No. 9.1e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 60

Qy     62 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 120

Qy    122 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttacaa 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttacaa 180

Qy    182 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 240

Qy    242 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 300

Qy    302 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 360

Qy    362 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 421
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 420

Qy    422 tctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattccc 481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 tctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattccc 480
```

```
Qy    482 ctgcta 487
          ||||||
Db    481 ctgcta 486


RESULT   3
AAH83992
ID   AAH83992 standard; DNA; 487 BP.
XX
AC   AAH83992;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Hylobates lar olfactory receptor encoding gene 4.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance; ds.
XX
OS   Hylobates lar.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   P-PSDB; AAG98506.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 1; Page 238; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence 487 BP; 129 A; 107 C; 78 G; 173 T; 0 other;


  Query Match             98.0%;  Score 477.4;  DB 22;  Length 487;
  Best Local Similarity   98.8%;  Pred. No. 1.5e-113;
  Matches  481;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gttgctaagcatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaatgatatttatttttgg 240

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db    241 tgcttttatacaaatacccactttaatgacgatcataatctcttattctcgtgtgctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcagcgc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    421 atctggcttagctgaagacccagacaaagtgtattctctgttttacacgattataattcc 480

Qy    481 cctgcta 487
          |||||||
Db    481 cctgcta 487


RESULT   4
AAH83986
ID   AAH83986 standard; DNA; 485 BP.
XX
AC   AAH83986;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Pan troglodytes olfactory receptor encoding gene 10.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance; ds.
XX
OS   Pan troglodytes.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   P-PSDB; AAG98501.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 1; Page 227; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence 485 BP; 130 A; 107 C; 77 G; 171 T; 0 other;


  Query Match             95.5%;  Score 465;  DB 22;  Length 485;
  Best Local Similarity   98.6%;  Pred. No. 2.3e-110;
  Matches  480;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    1;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
```

```
Db      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db     61 gttgctaagcatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
Db    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaatgatatttatttttgg 240

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db    241 tgcttttatacaaatacccactttaatgacgatcataatctcttatactcgtgtgctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcagcgc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db    421 atctggcttagctgaagacccagacaaagtgtattctc--ttttacacgattataattcc 478

Qy    481 cctgcta 487
Db    479 cctgcta 485
```

RESULT 5
AAH32168
ID AAH32168 standard; DNA; 988 BP.
XX
AC AAH32168;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 741.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 478; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 988 BP; 250 A; 231 C; 165 G; 342 T; 0 other;

Query Match 65.5%; Score 319; DB 22; Length 988;
Best Local Similarity 78.4%; Pred. No. 1.1e-72;
Matches 382; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

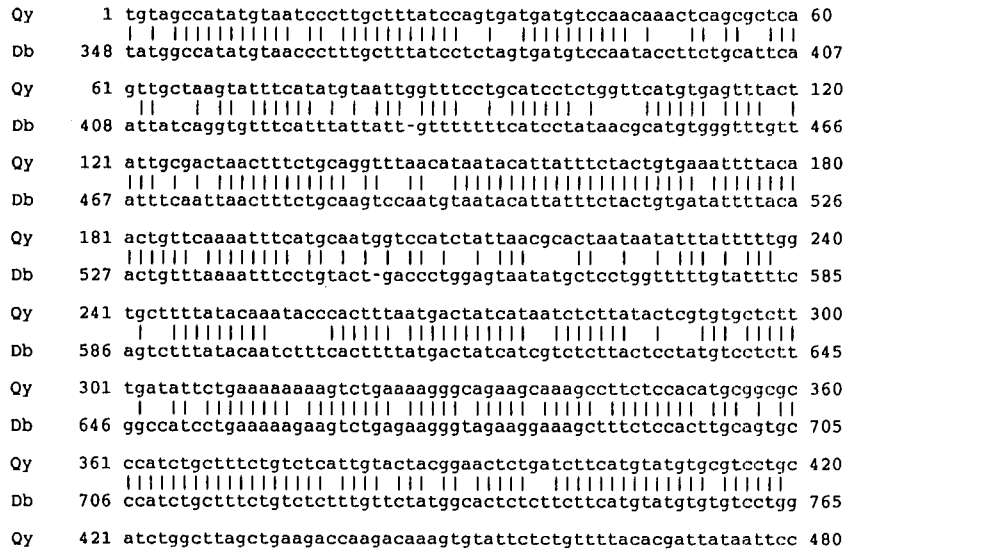
```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db    370 tgtagccatatgcaatcccttgctttatccagtggtgatgtccaatagcctctgtactca 429

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db    430 gtttataggtatttcatattttattggtttctgcattcagcgattcatgtgggtttgtt 489

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
Db    490 atttagattaactttctgcaggtccaatattatacattatttctactgtgaaatttttaca 549

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db    550 gctgttcaaaatttcttgcaccaatcctacagttaatatacttctgattttcatcttttc 609

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db    610 agcatttatacaagtcttcacttttatgactcttatcgtctcttactcctatattctctc 669

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db    670 tgccatcctgaaaaagaagtctgagaagggtagaagcaaagccttctctacttgcagtgc 729

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db    730 ccatctgctctctgtctctttgttctacggcaccctcttcttcatgtatgtgagttctag 789

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db    790 gtctggatcagctgcagatcaggccaaaatgtattctttatttacacaataataattcc 849

Qy    481 cctgcta 487
Db    850 tttacta 856
```

RESULT 6
AAI99556
ID AAI99556 standard; cDNA; 953 BP.
XX
AC AAI99556;
XX
DT 04-JAN-2002 (first entry)
XX
DE Human expressed polynucleotide SEQ ID NO 19.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;

KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200155387-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01310.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR P-PSDB; AAM99944.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/ or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 1; SEQ ID NO 19; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 953 BP; 242 A; 191 C; 181 G; 339 T; 0 other;

Query Match 56.6%; Score 275.8; DB 22; Length 953;
Best Local Similarity 75.6%; Pred. No. 1.5e-61;
Matches 368; Conservative 0; Mismatches 117; Indels 2; Gaps 2;



Qy 1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db 348 tatggccatatgtaaccctttgctttatcctctagtgatgtccaatacctttctgcattca 407

Qy 61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db 408 attatcaggtgtttcatttattatt-gttttttttcatcctataacgcatgtgggtttgtt 466

Qy 121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
Db 467 atttcaattaactttctgcaagtccaatgtaatacattatttctactgtgatatttttaca 526

Qy 181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db 527 actgtttaaaatttcctgtact-gaccctggagtaatatgctcctggtttttgtattttc 585

Qy 241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
Db 586 agtctttatacaatctttcacttttatgactatcatcgtctcttactcctatgtcctctt 645

Qy 301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db 646 ggccatcctgaaaaagaagtctgagaagggtagaaggaaagctttctccacttgcagtgc 705

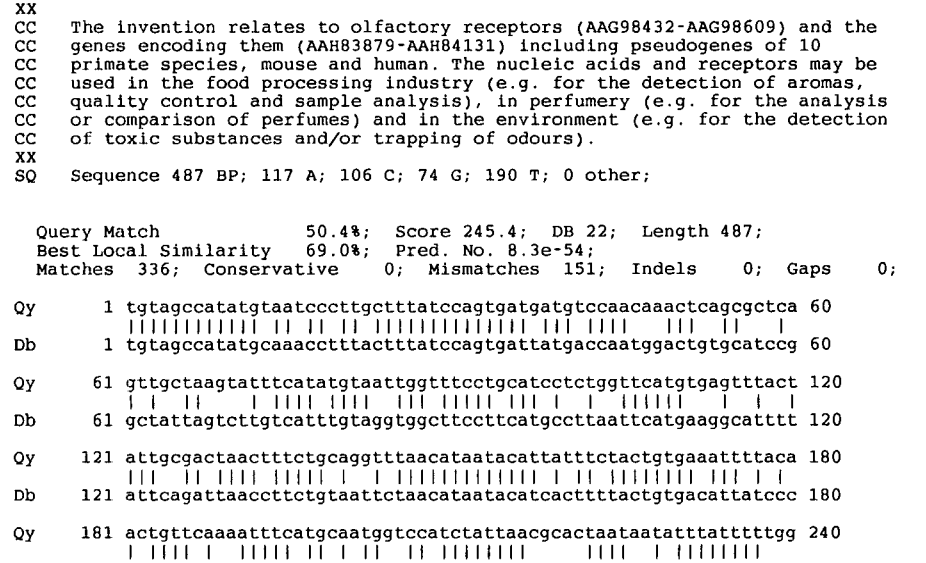
Qy 361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db 706 ccatctgctttctgtctctttgttctatggcactctcttcttcatgtatgtgtgtcctgg 765

Qy 421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db 766 gtctggatcagctgcagatcaggacaaaatatattctttatttttacatgataataattcc 825

Qy 481 cctgcta 487
Db 826 tctggta 832

RESULT 7
AAH83978
ID AAH83978 standard; DNA; 487 BP.
XX
AC AAH83978;
XX
DT 25-SEP-2001 (first entry)
XX
DE Pan troglodytes olfactory receptor encoding gene 3.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance; ds.
XX
OS Pan troglodytes.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB02017.
XX
PR 22-DEC-1999; 99US-0171746.
PR 21-DEC-2000; 2000US-0747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR P-PSDB; AAG98494.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful
PT for analysis odours e.g. in food processing and perfumery -
XX
PS Claim 1; Page 212-213; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 487 BP; 117 A; 106 C; 74 G; 190 T; 0 other;

Query Match 50.4%; Score 245.4; DB 22; Length 487;
Best Local Similarity 69.0%; Pred. No. 8.3e-54;
Matches 336; Conservative 0; Mismatches 151; Indels 0; Gaps 0;



Qy 1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db 1 tgtagccatatgcaaacctttacttttatccagtgattatgaccaatggactgtgcatccg 60

Qy 61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db 61 gctattagtcttgtcatttgtaggtggcttccttcatgccttaattcatgaaggcatttt 120

Qy 121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
Db 121 attcagattaaccttctgtaattctaacataatacatcactttactgtgacattatccc 180

Qy 181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240

```
Db    181 attgttaacgatttcctgtactgacccttctattaatttttaatgctttttattttgtc 240

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          || ||  ||||| ||  |||| |  ||||| |  |  |||||||| |   ||| |||||
Db    241 tggttcaatacaggtattcactattttgactgttcttgtctcttatgcatttgtcctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |   ||  | ||||||||||| |  || |||| ||| |||||||| ||||| || || ||
Db    301 tacaatcttaaaaaaaaagtcagtcaaaggcataaggaaagccttttccacctgtggagc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||  | |||||||  || ||||| ||  | ||  ||||||||||||||| | |||||
Db    361 ccatctcttctctgtctgtttatactatggccccttctcttcatgtatgtgggccctgc 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||     ||| || || ||||| |  ||  | | | | ||||||||  | || |||||
Db    421 atctccacaagcagatgatcaagatatggtagagtgtgtattttacactgtcatcattcc 480

Qy    481 cctgcta 487
            |  ||
Db    481 tttctta 487


RESULT   8
AAH31676
ID   AAH31676 standard; DNA; 942 BP.
XX
AC   AAH31676;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 249.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 8; Page 276-277; 1857pp; English.
XX
CC   The present sequence is one of a number of isolated polynucleotides
CC   which encode polypeptides involved in olfactory sensation. The
CC   polynucleotides can be used in screening for olfactory agonists and
CC   antagonists. The methods allow for the determination of primary
CC   scents and the identification of the odour receptors used to detect
CC   these primary scents. The methods also enable determination of
CC   secondary scents and the identification of combinations of odour
CC   receptors that are involved in detecting such secondary scents.
CC   This enables the construction of a scent representation (also called
CC   a scent fingerprint or scent profile), which may be used to re-create
CC   and edit scents. Libraries of olfactory receptors are useful for
CC   determining the interaction pattern of a composition with the receptors,
CC   and can be used for determining differences in the olfactory faculties
CC   of different individuals.
XX
SQ   Sequence 942 BP; 255 A; 196 C; 152 G; 339 T; 0 other;


  Query Match             49.7%;  Score 242.2;  DB 22;  Length 942;
  Best Local Similarity   68.6%;  Pred. No. 6.6e-53;
  Matches  334;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          |||||||||||| || || || || |||||||||||| |||  |||   |||  ||   |
Db    384 tgtagccatatgcaaacctttactatatccagtgattatgaacaattcactatgcatacg 443

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | || ||    | |||| | ||  ||| ||||| ||| |  |  ||||||      | |
Db    444 gctgttagccttctcatttttaggtggcttcctccatgccttaattcatgaagtccttat 503

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
          |||  || |||| |||||||  | |||||||||||||| |||| |||||||| ||| |||
Db    504 attcagattaaccttctgcaattctaacataatacatcatttttactgtgatattatacc 563

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||| |  ||||| || | ||  || |||||||     |||||  | |||||||||
Db    564 actgtttatgatttcctgtactgacccttctattaattttctaatggtttttatttgtc 623

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          ||  |  || ||  ||  ||| |  |||| |  | |  ||||| ||   || .|| ||
Db    624 tggctcaattcaggtattcaccattgtgacagttcttaattcttacacatttgctctttt 683

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
              || || ||||| |||||||  |  |||  ||| |||||||| ||||| || || ||
Db    684 cacaatcctaaaaaagaagtctgttagaggcgtaaggaaagccttttccacctgtggagc 743

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||  | |||||||| || || || ||  | || ||||||||||||| |||| |||||
Db    744 ccatctcttatctgtctctttatattatggcccacttatcttcatgtatttgcgccctgc 803

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||     ||| || |||||||| |   |  | ||| | ||||| || || ||||||||
Db    804 atctccacaagcagatgaccaagatatgatagactctgtcttttatacaatcataattcc 863

Qy    481 cctgcta 487
            |||||
Db    864 tttgcta 870


RESULT   9
AAS42326
ID   AAS42326 standard; cDNA; 945 BP.
XX
AC   AAS42326;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human cDNA encoding olfactory receptor AOLFR127.
XX
KW   Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW   ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS   Homo sapiens.
XX
PN   WO200168805-A2.
XX
PD   20-SEP-2001.
XX
PF   13-MAR-2001; 2001WO-US07771.
XX
PR   13-MAR-2000; 2000US-0188914.
PR   24-MAR-2000; 2000US-0192033.
PR   12-APR-2000; 2000US-0198474.
PR   24-APR-2000; 2000US-0199335.
PR   26-MAY-2000; 2000US-0207702.
PR   23-JUN-2000; 2000US-0213849.
```

PR 16-AUG-2000; 2000US-0226534.
PR 07-SEP-2000; 2000US-0230732.
PR 07-FEB-2001; 2001US-0266862.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR P-PSDB; AAU24633.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors,
PT useful for screening for compounds involved in olfactory sensation,
PT where the compounds can be used in the food, pharmaceutical and
PT cosmetic industries to customise odours -
XX
PS Claim 1; Page 136; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence encodes a human olfactory receptor of
CC the invention.
XX
SQ Sequence 945 BP; 256 A; 196 C; 153 G; 340 T; 0 other;

Query Match 49.7%; Score 242.2; DB 22; Length 945;
Best Local Similarity 68.6%; Pred. No. 6.6e-53;
Matches 334; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||| || || || || |||||||||||| |||  |||   |||  ||   |
Db    384 tgtagccatatgcaaaccttactatatccagtgattatgaacaattcactatgcatacg 443

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | || ||    | |||| | ||  ||| ||||| ||| |  |  ||||||      |  |
Db    444 gctgttagccttctcatttttaggtggcttcctccatgccttaattcatgaagtccttat 503

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          |||  || |||| |||||||  | ||||||||||||| |||| |||||||| ||| |||
Db    504 attcagattaaccttctgcaattctaacataatacatcatttttactgtgatattatacc 563

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||| |  ||||| || | ||  || ||||||||    |||||  | ||||||||
Db    564 actgtttatgatttcctgtactgacccttctattaattttctaatggtttttatttttgtc 623

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          ||  |  || ||  ||  |||  |  ||||  |  | |  ||||| ||   ||  || ||
Db    624 tggctcaattcaggtattcaccattgtgacagttcttaattcttacacatttgctctttt 683

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
             || ||||| |||||||  |  |||  ||| |||||||| ||||| || || ||
Db    684 cacaatcctaaaaaagaagtctgttagaggcgtaaggaaagccttttccacctgtggagc 743

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||  | |||||||| || || || ||  | || ||||||||||| |||| |||||
Db    744 ccatctcttatctgtctctttatattatggcccacttatcttcatgtatttgcgccctgc 803

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||     ||| || |||||||| |   |  | ||| | ||||| || || ||||||||
Db    804 atctccacaagcagatgaccaagatatgatagactctgtcttttatacaatcataattcc 863

Qy    481 cctgcta 487
            |||||
Db    864 tttgcta 870
```

RESULT 10
AAH31707
ID AAH31707 standard; DNA; 958 BP.
XX
AC AAH31707;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 280.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 289-290; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 958 BP; 261 A; 200 C; 150 G; 347 T; 0 other;

Query Match 48.5%; Score 236; DB 22; Length 958;
Best Local Similarity 69.1%; Pred. No. 2.6e-51;
Matches 337; Conservative 0; Mismatches 150; Indels 1; Gaps 1;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          |||||||||||| || || || |||||||||||| || ||| ||||   |||  |
Db    369 tgtagccatatgcaaaccttactttatccagtcattatgaccaatggactgtgtatctg 428

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | |  ||    | |||| | ||  ||| || || ||| |  |  ||||||   || |  |
Db    429 gctattagtcttgtcatttataggtggctttcttcatgccttaattcatgaaggtatttt 488

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          |||  || |||| ||||| |  |  ||||||||||| | || ||||||| ||| | |
Db    489 attcagattaaccttctgtaattccaacataatacatcacttttactgtgacattatccc 548

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
```

```
Db    549 attgttaaagatttcctgtactgacccttctattaatttttaatgctttttatttgtc 608

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db    609 tggttcaatacaggtattcactattttgactgttcttgtctcttatacatttgtcctctt 668

Qy    301 tgatat-tctgaaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcg 359
Db    669 tacaatcttaaaaaaaaaagtctgccaaagacataaggaaagccttttccacctgtggag 728

Qy    360 cccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctg 419
Db    729 cccatctcttatctgtttctttatactatggccccttctcttcatgtatgtgcaccctg 788

Qy    420 catctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattc 479
Db    789 catctccacaagcagatgatcaagatatggtggagtctctattttacactgtcataattc 848

Qy    480 ccctgcta 487
Db    849 ctttctta 856
```

RESULT 11
AAH31705
ID AAH31705 standard; DNA; 972 BP.
XX
AC AAH31705;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 278.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 288-289; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 972 BP; 269 A; 203 C; 151 G; 349 T; 0 other;

```
  Query Match             47.1%;  Score 229.4;  DB 22;  Length 972;
  Best Local Similarity   66.9%;  Pred. No. 1.3e-49;
  Matches  326;  Conservative    0;  Mismatches  161;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db    369 tgtagccatttgcaaagctttactttatccagtcattatgaccaacggactatgcatcca 428

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db    429 gctattagtcttgtcatttataggtggccttcttcatgctttaatccatgaaatttttt 488

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttaca 180
Db    489 attcagattaaccttctgtaattccaacataatacaacacttttactgtgacattatccc 548

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db    549 attgttaaagatttcctgtactgattcttttattaactttctaatggtttttatttcgc 608

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db    609 agattcaattcaagtttttaccattggaactattcttatatcttatacacttgtcctcct 668

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db    669 tataatcttaaaaaataagtctgtcaaagggatacaaaaagctgtctccacctgtggagc 728

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db    729 tcatctcttatctgtatctttatactatgggccccttgtcttcatgtatgtgggctctgc 788

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db    789 atccccgcaagcagatgaccaagatatgatggagtctctattttacactgtcatcgttcc 848

Qy    481 cctgcta 487
Db    849 tttatta 855
```

RESULT 12
AAH31701
ID AAH31701 standard; DNA; 927 BP.
XX
AC AAH31701;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 274.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 287; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 927 BP; 256 A; 203 C; 143 G; 325 T; 0 other;

Query Match 46.1%; Score 224.6; DB 22; Length 927;
Best Local Similarity 66.3%; Pred. No. 2.2e-48;
Matches 323; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

```
Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db     369 tgtagccatttgcaaagctttacttttatccagtcattatgaccaatgaactatgcattca 428

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db     429 gctattagtcttgtcatttataggtggccttcttcatgctttaatccatgaagctttttc 488

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
Db     489 attcagattaaccttctgtaattccaacataatacaacactttactgtgacattatccc 548

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db     549 attgttaaagatttcctgtactgattcctctattaactttctaatggtttttatttcgc 608

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db     609 aggttctgttcaagtttttaccattggaactattcttatatcttatacaattatcctctt 668

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db     669 tacaatcttagaaaagaagtctatcaaagggatacgaaaagctgtctccacctgtggggc 728

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db     729 tcatctcttatctgtatctttatactatggccccctcaccttcaaatatctgggctctgc 788

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db     789 atctccgcaagcagatgaccaagatatgatggagtctctatttacactgtcatagttcc 848

Qy     481 cctgcta 487
Db     849 tttatta 855
```

RESULT 13
AAS42325
ID AAS42325 standard; cDNA; 978 BP.
XX
AC AAS42325;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human cDNA encoding olfactory receptor AOLFR126.
XX
KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS Homo sapiens.
XX
PN WO200168805-A2.
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07771.
XX
PR 13-MAR-2000; 2000US-0188914.
PR 24-MAR-2000; 2000US-0192033.
PR 12-APR-2000; 2000US-0198474.
PR 24-APR-2000; 2000US-0199335.
PR 26-MAY-2000; 2000US-0207702.
PR 23-JUN-2000; 2000US-0213849.
PR 16-AUG-2000; 2000US-0226534.
PR 07-SEP-2000; 2000US-0230732.
PR 07-FEB-2001; 2001US-0266862.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR P-PSDB; AAU24632.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors,
PT useful for screening for compounds involved in olfactory sensation,
PT where the compounds can be used in the food, pharmaceutical and
PT cosmetic industries to customise odours -
XX
PS Claim 1; Page 136; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence encodes a human olfactory receptor of
CC the invention.
XX
SQ Sequence 978 BP; 267 A; 212 C; 155 G; 344 T; 0 other;

Query Match 46.1%; Score 224.6; DB 22; Length 978;
Best Local Similarity 66.3%; Pred. No. 2.3e-48;
Matches 323; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

```
Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db     417 tgtagccatttgcaaagctttacttttatccagtcattatgaccaatgaactatgcattca 476

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db     477 gctattagtcttgtcatttataggtggccttcttcatgctttaatccatgaagctttttc 536

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
Db     537 attcagattaaccttctgtaattccaacataatacaacactttactgtgacattatccc 596
```

```
Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          | |||| || ||||| || | || | | |||||||||   |||||  | |||||||| |
Db    597 attgttaaagatttcctgtactgattcctctattaactttctaatggtttttattttcgc 656

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           | || | | ||| |    ||  |    |||||  | || ||||||||   | | |||||
Db    657 aggttctgttcaagtttttaccattggaactattcttatatcttatacaattatcctctt 716

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |   ||  |  |||| ||||||   || || | | | |||||  ||||||| || || ||
Db    717 tacaatcttagaaaagaagtctatcaaagggatacgaaaagctgtctccacctgtggggc 776

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           |||||  | ||||| || || ||||| ||  | || | |||||  ||| || |  ||||
Db    777 tcatctcttatctgtatctttatactatggccccctcaccttcaaatatctgggctctgc 836

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||     ||| || |||||||| |   || | ||||| |||||||| | ||| ||||
Db    837 atctccgcaagcagatgaccaagatatgatggagtctctattttacactgtcatagttcc 896

Qy    481 cctgcta 487
            |  ||
Db    897 tttatta 903


RESULT 14
AAH31698
ID   AAH31698 standard; DNA; 940 BP.
XX
AC   AAH31698;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 271.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 8; Page 285-286; 1857pp; English.
XX
CC   The present sequence is one of a number of isolated polynucleotides
CC   which encode polypeptides involved in olfactory sensation. The
CC   polynucleotides can be used in screening for olfactory agonists and
CC   antagonists. The methods allow for the determination of primary
CC   scents and the identification of the odour receptors used to detect
CC   these primary scents. The methods also enable determination of
CC   secondary scents and the identification of combinations of odour
CC   receptors that are involved in detecting such secondary scents.
CC   This enables the construction of a scent representation (also called
CC   a scent fingerprint or scent profile), which may be used to re-create
CC   and edit scents. Libraries of olfactory receptors are useful for
CC   determining the interaction pattern of a composition with the receptors,
CC   and can be used for determining differences in the olfactory faculties
CC   of different individuals.
XX
SQ   Sequence 940 BP; 258 A; 192 C; 146 G; 344 T; 0 other;


  Query Match             44.6%;  Score 217;  DB 22;  Length 940;
  Best Local Similarity   66.7%;  Pred. No. 2e-46;
  Matches  325;  Conservative    0;  Mismatches  160;  Indels    2;  Gaps    1;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          |||||||||||| || || || |||||||||||| || ||| ||||   |||  ||
Db    366 tgtagccatatgcaaaccttacttttatccagtcattatgaccaatggactgtgcatctg 425

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | |  |||   | |||| | ||  |||  | || ||| || |  ||||||   ||||  |
Db    426 gctattaatcttgtcatttctaggtggccttcttcatgctttaattcatgaaggttttt 485

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||   || |||| ||||| |  |  |||| ||||||| | || || ||||| ||| | |
Db    486 at--agattaaccttctgtaattccaacacaatacatcacttttaatgtgacattatccc 543

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          | |||| |||||||  || | || | | |||||||||   | |||  | ||||||||
Db    544 attgttaaaaattttctgtactgattcttctattaactttccaatggtttttatttctc 603

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
             ||  || ||| |   |||  |   |||| |  |  | ||||||||   ||| ||||
Db    604 atgttcaattcaagttttcaccattgggactgttcttgtatcttatacatttgtcctctc 663

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |   ||  ||||||| |||||||  || |||| ||| |||| |||||||||| || || ||
Db    664 tacaatcttgaaaaagaagtctgtcaaaggcataagaaaagacttctccacctgtggagc 723

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||| |  |  |||| || || ||||| ||  | ||   ||||||||||||| |  ||||
Db    724 tcatatcttacctgtatctttatactatgggcccctcgccttcatgtatgtgggctctgc 783

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||       ||||| |||||||| |   || | ||||| ||||||||  | ||| ||||
Db    784 atcccaacgggctgatgaccaagatatgatggagtctctattttacactgtcatagttcc 843

Qy    481 cctgcta 487
            |  ||
Db    844 tttatta 850


RESULT 15
AAH31696
ID   AAH31696 standard; DNA; 944 BP.
XX
AC   AAH31696;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 269.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
```

PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 285; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 944 BP; 262 A; 184 C; 152 G; 346 T; 0 other;

Query Match 43.2%; Score 210.2; DB 22; Length 944;
Best Local Similarity 64.5%; Pred. No. 1.1e-44;
Matches 314; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          || || |||||| || || || |||||||||||| || ||| |||| |||| ||
Db    376 tgcagacatatgcaaaccttactttatccagtcattatgaccaatgaactatgcatctg 435

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | | |     | |||| | ||  |||  |  | ||| || |  | ||||   ||||  |
Db    436 gctatttgtcttgtcatttctaggtggccttttcatgctttaatccatgaaggtttttt 495

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          |||  ||||||| ||||| |  |  ||||| |||||  |||| |||||||| ||| | |
Db    496 attcagactaaccttctgtaactccaacatgatacaacatttttactgtgacattatccc 555

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          | |||| || |||||||| | || | | | ||||||    |||||  | ||||||||
Db    556 attgttaaagatttcatgtactgattcttgtattaattttctaatgttttttatttctc 615

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           | ||  || ||| |    ||  |   || | |  |  | |||||||    ||| |||||
Db    616 aggttcaattcaagttttaaccattgggattgtttttgtatcttatatgtttgttctctt 675

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |   ||  | ||||| |||||| | || |||| ||| |||||| ||||| || || ||
Db    676 tacaatcttaaaaaagaagtctaacaaaggcataagggaagccttttccacctgtggagc 735

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||    |  || |||| || |  || ||    ||  |||||||||||||| | |||||
Db    736 ccattacatacctctctctttatgttatggcctccttctcttcatgtatgtgggccctgc 795

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          | ||     ||| ||  | ||||| |   || | | ||| ||||||  | || ||  | |
Db    796 agctccacaagcagataatcaagatatgatggagtatctattttacccctatcattgtgcg 855

Qy    481 cctgcta 487
            || ||
Db    856 tttgtta 862
```

Search completed: June 27, 2002, 11:20:53

Job time: 3746 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 10:18:27 ; Search time 43.34 Seconds
(without alignments)
2760.118 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	103	21.1	1713	2	US-08-467-948A-1	Sequence 1, Appli
	2	103	21.1	1713	3	US-08-467-947A-1	Sequence 1, Appli
	3	93.2	19.1	1290	2	US-08-827-291A-1	Sequence 1, Appli
	4	90.8	18.6	966	3	US-08-748-506-8	Sequence 8, Appli
	5	88.4	18.2	966	3	US-08-748-506-6	Sequence 6, Appli
	6	85.2	17.5	966	3	US-08-748-506-5	Sequence 5, Appli
	7	84.4	17.3	900	4	US-09-085-371-5	Sequence 5, Appli
	8	73.2	15.0	966	3	US-08-748-506-7	Sequence 7, Appli
	9	71.6	14.7	984	3	US-08-748-506-9	Sequence 9, Appli
c	10	51.4	10.6	1320	1	US-08-599-252-84	Sequence 84, Appl
c	11	51.4	10.6	1320	1	US-08-436-074-57	Sequence 57, Appl
c	12	51.4	10.6	1320	5	PCT-US96-06352-84	Sequence 84, Appl
c	13	51.4	10.6	1320	5	PCT-US96-06583-84	Sequence 84, Appl
	14	45.4	9.3	7218	1	US-08-232-463-14	Sequence 14, Appl
c	15	38.4	7.9	2251	4	US-08-991-677-11	Sequence 11, Appl
	16	37.2	7.6	2429	4	US-09-386-493-3	Sequence 3, Appli
	17	34.6	7.1	236	2	US-08-744-905A-1	Sequence 1, Appli
	18	34.6	7.1	1241	1	US-07-593-657-6	Sequence 6, Appli
	19	34.6	7.1	1241	4	US-08-942-012B-3	Sequence 3, Appli
	20	34.4	7.1	6243	2	US-09-056-075-1	Sequence 1, Appli
	21	34	7.0	1828	3	US-08-988-876-2	Sequence 2, Appli
c	22	33.4	6.9	10968	2	US-08-680-327-2	Sequence 2, Appli
c	23	33.4	6.9	10968	4	US-09-228-246-1	Sequence 1, Appli
c	24	32.8	6.7	246240	2	US-08-724-394A-20	Sequence 20, Appl
c	25	32.8	6.7	246240	2	US-08-724-394A-21	Sequence 21, Appl
c	26	32.8	6.7	246240	2	US-08-724-394A-22	Sequence 22, Appl
	27	32.6	6.7	565	2	US-08-731-722-7	Sequence 7, Appli
	28	32.6	6.7	934	2	US-08-731-722-6	Sequence 6, Appli
c	29	32.6	6.7	934	2	US-08-731-722-6	Sequence 6, Appli
	30	32.6	6.7	1183	2	US-08-731-722-8	Sequence 8, Appli
c	31	32.6	6.7	1183	2	US-08-731-722-8	Sequence 8, Appli
	32	32.6	6.7	1186	2	US-08-731-722-5	Sequence 5, Appli
c	33	32.6	6.7	1186	2	US-08-731-722-5	Sequence 5, Appli
	34	32.6	6.7	1218	2	US-08-731-722-4	Sequence 4, Appli
c	35	32.6	6.7	1218	2	US-08-731-722-4	Sequence 4, Appli
	36	32.6	6.7	3926	2	US-08-731-722-1	Sequence 1, Appli
c	37	32.6	6.7	3926	2	US-08-731-722-1	Sequence 1, Appli
	38	32.6	6.7	3926	2	US-08-731-722-2	Sequence 2, Appli
c	39	32.6	6.7	3926	2	US-08-731-722-2	Sequence 2, Appli
	40	32.6	6.7	3933	2	US-08-731-722-3	Sequence 3, Appli
c	41	32.6	6.7	3933	2	US-08-731-722-3	Sequence 3, Appli
	42	32.4	6.7	15766	4	US-09-338-907-73	Sequence 73, Appl
	43	32.4	6.7	15766	4	US-09-218-207-73	Sequence 73, Appl
	44	32.4	6.7	37950	4	US-09-338-907-183	Sequence 183, App
	45	32.4	6.7	37950	4	US-09-218-207-183	Sequence 183, App

ALIGNMENTS

```
RESULT   1
US-08-467-948A-1
; Sequence 1, Application US/08467948A
; Patent No. 5998164
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR2
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,948A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04079
;      FILING DATE:  30-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1140003/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
```

```
;      LOCATION:  116..1003
US-08-467-948A-1


  Query Match             21.1%;  Score 103;  DB 2;  Length 1713;
  Best Local Similarity   51.4%;  Pred. No. 5.5e-20;
  Matches  238;  Conservative    0;  Mismatches  225;  Indels    0;  Gaps    0;

Qy      2 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 61
          || ||||| ||  | ||  | |  |||    | || ||| ||   ||| || ||
Db    494 GTGGCCATCTGCCACCCTCTCCGATATTTCATCATCATGACCTGGAAAGTCTGCATCACT 553

Qy     62 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 121
           ||    |  | ||| |    |  ||| | ||| |   || |||| ||||||||  |  |
Db    554 CTGGGCATCACTTCCTGGACATGTGGCTCCCTCCTGGCTATGGTCCATGTGAGCCTCATC 613

Qy    122 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttacaa 181
           |  ||||  | || ||  ||  |     |||  |  | |||| ||||||||  |
Db    614 CTAAGACTGCCCTTTTGTGGGCCTCGTGAAATCAACCACTTCTTCTGTGAAATCCTGTCT 673

Qy    182 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 241
           |  |||   |  | ||   || | | |   | |||    |  | || |||        |
Db    674 GTCCTCAGGCTGGCCTGTGCTGATACCTGGCTCAACCAGGTGGTCATCTTTGAAGCCTGC 733

Qy    242 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 301
          || || |   |     |  |       |  |  |||| ||  | |   |
Db    734 ATGTTCATCCTGGTGGGACCACTCTGCCTGGTGCTGGTCTCCTACTCACACATCCTGGGG 793

Qy    302 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 361
          |  || ||||  |  ||||||  ||||| | || || ||||||||||| |||  | ||
Db    794 GGCATCCTGAGGATCCAGTCTGGGGAGGGCCGCAGAAAGGCCTTCTCCACCTGCTCCTCC 853

Qy    362 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 421
          || ||       ||   | | | ||  || |     ||| |||||  ||   |||
Db    854 CACCTCTGCGTAGTGGGACTCTTCTTTGGSAGCGCCATCGTCATGTCATGGCCCCTAAG 913

Qy    422 tctggcttagctgaagaccaagacaaagtgtattctctgtttt 464
          ||  ||   |||| || ||  | || ||   || |||  |||
Db    914 TCCCGCCATCCTGAGGAGCAGCAGAAGGTCCTTTTTCTTATTT 956


RESULT   2
US-08-467-947A-1
; Sequence 1, Application US/08467947A
; Patent No. 6090575
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR1
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,947A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04079
;      FILING DATE:  30-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1140002/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  116..1003
US-08-467-947A-1


  Query Match             21.1%;  Score 103;  DB 3;  Length 1713;
  Best Local Similarity   51.4%;  Pred. No. 5.5e-20;
  Matches  238;  Conservative    0;  Mismatches  225;  Indels    0;  Gaps    0;

Qy      2 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 61
          || ||||| ||  | ||  | |  |||    | || ||| ||   ||| || ||
Db    494 GTGGCCATCTGCCACCCTCTCCGATATTTCATCATCATGACCTGGAAAGTCTGCATCACT 553

Qy     62 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 121
           ||    |  | ||| |    |  ||| | ||| |   || |||| ||||||||  |  |
Db    554 CTGGGCATCACTTCCTGGACATGTGGCTCCCTCCTGGCTATGGTCCATGTGAGCCTCATC 613

Qy    122 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttacaa 181
           |  ||||  | || ||  ||  |     |||  |  | |||| ||||||||  |
Db    614 CTAAGACTGCCCTTTTGTGGGCCTCGTGAAATCAACCACTTCTTCTGTGAAATCCTGTCT 673

Qy    182 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 241
           |  |||   |  | ||   || | | |   | |||    |  | || |||        |
Db    674 GTCCTCAGGCTGGCCTGTGCTGATACCTGGCTCAACCAGGTGGTCATCTTTGAAGCCTGC 733

Qy    242 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 301
          || || |   |     |  |       |  |  |||| ||  | |   |
Db    734 ATGTTCATCCTGGTGGGACCACTCTGCCTGGTGCTGGTCTCCTACTCACACATCCTGGGG 793

Qy    302 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 361
          |  || ||||  |  ||||||  ||||| | || || ||||||||||| |||  | ||
Db    794 GGCATCCTGAGGATCCAGTCTGGGGAGGGCCGCAGAAAGGCCTTCTCCACCTGCTCCTCC 853

Qy    362 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 421
          || ||       ||   | | | ||  || |     ||| |||||  ||   |||
Db    854 CACCTCTGCGTAGTGGGACTCTTCTTTGGSAGCGCCATCGTCATGTACATGGCCCCTAAG 913

Qy    422 tctggcttagctgaagaccaagacaaagtgtattctctgtttt 464
          ||  ||   |||| || ||  | || ||   || |||  |||
Db    914 TCCCGCCATCCTGAGGAGCAGCAGAAGGTCCTTTTTCTTATTT 956


RESULT   3
US-08-827-291A-1
; Sequence 1, Application US/08827291A
; Patent No. 5874243
;  GENERAL INFORMATION:
;    APPLICANT:  Macina, Roberto
;    APPLICANT:  Sathe, Ganesh
;    TITLE OF INVENTION:  NOVEL OLRCC15 RECEPTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
```

```
;       COUNTRY:
;       ZIP:  19406
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/827,291A
;       FILING DATE:  28-MAR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  King, William T
;       REGISTRATION NUMBER:  30,954
;       REFERENCE/DOCKET NUMBER:  GP50001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610-270-5015
;       TELEFAX:  610-270-5090
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1290 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-827-291A-1


  Query Match             19.1%;  Score 93.2;  DB 2;  Length 1290;
  Best Local Similarity   49.6%;  Pred. No. 2.8e-17;
  Matches  239;  Conservative    0;  Mismatches  243;  Indels    0;  Gaps    0;

Qy       5 gccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcagttg 64
Db     668 GCCATTTGCCACCCTCTAAGATACACCAATCTCATGAGCCCTAAAATTTGTGGACTTATG 727

Qy      65 ctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactattg 124
Db     728 ACTGCCTTTTCCTGGATCCTGGGCTCTACAGATGGAATCATTTATGCTGTAGCCACATTT 787

Qy     125 cgactaactttctgcaggtttaacataatacattatttctactgtgaaaatttacaactg 184
Db     788 TCCTTCTCCTACTGTGGGTCTCGGGAAATAGCCCACTTCTTCTGTGAGTTACCTTCCCTA 847

Qy     185 ttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggtgct 244
Db     848 CTAATCCTCTCATGCAATGACACATCAATATTTGAAAAGGTTATTTTCATTTGCTCTATA 907

Qy     245 tttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctttgat 304
Db     908 GTAATGCTTGTTTTCCCTGTTGCAATCATCATTGCTTCCTATGCTGGAGTTATTCTGGCT 967

Qy     305 attctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcccat 364
Db     968 GTCATTCACATGGGATCTGGAGAGGGTCGTCGCAAAGCTTTCACGACCTGTTCCTCTCAC 1027

Qy     365 ctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcatct 424
Db    1028 CTCATGGTGGTGGGAATGTTCTATGGAGCAGGTTTGTTCATGTACATACAGCCCACATCT 1087

Qy     425 ggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcccctg 484
Db    1088 GATCGCTCCCCAACGCAGGACAAGCTGGTGTCTGTATTCTACACCATCCTCACTCCCATG 1147

Qy     485 ct 486
Db    1148 CT 1149
```

```
RESULT   4
US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-8


  Query Match             18.6%;  Score 90.8;  DB 3;  Length 966;
  Best Local Similarity   49.2%;  Pred. No. 1.2e-16;
  Matches  239;  Conservative    0;  Mismatches  247;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db     387 TATGGCTATATGTTCCCCACTCCACTATGCAACCCGAATGAGTCGTGAGGTATGTGCCCA 446

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db     447 CTTGGCAATTGTTTCATGGGTGATGGGATGCATAGTAGGTCTGGGACAGACCAATTTTAT 506

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaaatttaca 180
Db     507 TTTCTCCTTGAACTTCTGTGGACCCTGTGAGATAGACCACTTCTTCTGTGATCTTCCACC 566

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db     567 TCTCCTGGCACTTGCCTGTGGTGATACATCCCAAATTGAGGCTGCCATCTTTGTGGTAGT 626

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db     627 TGTCCTCTGCATATCTAGCCCTTTTTTGCTGATCATTTATTCTTATGTCAGAATTCTCGT 686

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db     687 TGCAGTGCTGGTGATGCCTTCACCTGAGGGGCGCCACAAAGCCCTTTCAACCTGTTCCTC 746

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
```

```
Db    747 CCACCTACTTGTAGTCACACTCTTTTATGGCTCAGGATCTGTTACCTATTTGAGGCCTAA 806

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||| ||    |  ||     |||||| | |   | || || |||||     | |   |
Db    807 GTCTAGCCACTCACCAGGAATGGACAAACTCTTGGCCCTCTTCTACACAGCAGTGACATC 866

Qy    481 cctgct 486
          | || |
Db    867 CATGTT 872


RESULT   5
US-08-748-506-6
; Sequence 6, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-6


  Query Match             18.2%;  Score 88.4;  DB 3;  Length 966;
  Best Local Similarity   48.8%;  Pred. No. 5.7e-16;
  Matches  236;  Conservative    0;  Mismatches  248;  Indels    0;  Gaps    0;

Qy      3 tagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcagt 62
          | | ||||||    ||  | |  ||| ||      |||       |  |  | || || |
Db    389 TGGGCATATGCTCCCCACTCCACTATGCAACCCGAATGAGTCGTGAGGTATGTGCCCATT 448

Qy     63 tgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactat 122
          ||  || | ||||||  | ||| || | | |      |||||  ||    | |||  | |
Db    449 TGGCAATTGTTTCATGGGGAATGGGATGCATAGTAGGTCTGGGACAGACCAATTTNATTT 508

Qy    123 tgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttacaac 182
            | |  |||||  |   |     ||| |  | |||| ||||||  ||  ||  |
Db    509 NCTCCTTGAACTTCTGTGGACCTTGTGAGATAGACCACTTCTTCTGTGACCTTCCACCTC 568

Qy    183 tgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggtg 242
          |  |   | || | |   || | ||||    ||||        || ||| |    |
Db    569 TCCTGGCACTTGCCTGTGGTGATACATCCCAAAACGAGGCTGCCATCTTTGTGGCAGCAA 628

Qy    243 cttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctttg 302
          |      |       | ||  ||   ||| |    ||||||    |  | ||| |||
Db    629 TCCTCTGTATATCTAGTCCATTTTTGGTGATCCTTTATTCTTATGTCAGAATTCTCGTTG 688

Qy    303 atattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgccc 362
          | |||   |     ||     ||||  |   |||||  | ||||| ||   | | |
Db    689 CAGTGCTGGTGATGCCTTCACCTGAGGGGCGCCATAAAGCTCTTTCCACCTGTTCCTCAC 748

Qy    363 atctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcat 422
          | || |||   ||| || | |  || ||  || ||  ||| |  ||| || | ||    |
Db    749 ACCTACTTGTAGTCACACTCTTTTATGGCTCTGTGTCCTTTACCTATTTGAGGCCCAAGT 808

Qy    423 ctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcccc 482
          || ||    |  ||     ||||||  | |   | || || |||||     | |   ||
Db    809 CTAGCCACTCACCAGGAATGGACAAACTCTTGGCCCTCTTCTACACAGCAGTGACATCCA 868

Qy    483 tgct 486
          ||||
Db    869 TGCT 872


RESULT   6
US-08-748-506-5
; Sequence 5, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-5


  Query Match             17.5%;  Score 85.2;  DB 3;  Length 966;
  Best Local Similarity   48.3%;  Pred. No. 4.5e-15;
  Matches  234;  Conservative    0;  Mismatches  250;  Indels    0;  Gaps    0;

Qy      3 tagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcagt 62
          | ||||||||    ||  | |  ||| ||      |||       |  |  | || || |
Db    389 TGGCCATATGCTCCCCACTCCACTATGCAACCCGAATGAGTCGNGAGGTATGTGCCCATT 448
```

```
Qy      63 tgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactat 122
           ||  || | ||||||  | ||| || | | |      |||||  ||    | |||  | |
Db     449 TGGCAATTGTTTCATGGGGAATGGGATGCATAGTAAGTCTGGGACAAACCAATTTTATTT 508

Qy     123 tgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttacaac 182
           |    | |  |||||  |          |||| |  | |||| ||||||  ||  ||  |
Db     509 TCTCCTTGAACTTCTGTGGACCCTGTGAAATAGACCACTTCTTCTGTGACCTTCCACCTC 568

Qy     183 tgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggtg 242
           |  |   | || | ||    | | ||||    ||||        || ||| |  | |  |
Db     569 TCCTGGCACTTGCCTGTGGAGATACATCCCAAAACGAGGCTGCCATCTTTGTGGTAGCAG 628

Qy     243 cttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctttg 302
           |       |     | | ||  ||   |||||    ||||||       | ||| |||
Db     629 TCCTCTGCATATCTAGCCCATTTTTGCTGATCATTTATTCTTATGTCAAAATTCTCATTG 688

Qy     303 atattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgccc 362
           | ||    |     ||     ||||  |   |||||  | ||||| ||      | |
Db     689 CAGTGCTNCTGATGCCTTCACCTGAGGGGCGCCATAAAGCTCTTTCCACCTGTTCGTCTC 748

Qy     363 atctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcat 422
           | || |||   ||| || | |  || ||  |       | |  ||| || | ||    |
Db     749 ACCTACTTGTAGTCACACTTTTTTATGGCTCAGCATGTATTACCTATTTGAGGCCCAAGT 808

Qy     423 ctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcccc 482
           || ||    |   ||     |||||| | |   | || || |||||  |  | |   ||
Db     809 CTAGCCACTCACCAGGAATGGACAAATTCTTGGCCCTCTTCTACACAGTAGTGACATCCA 868

Qy     483 tgct 486
           ||||
Db     869 TGCT 872
```

RESULT 7
US-09-085-371-5
; Sequence 5, Application US/09085371
; Patent No. 6218358
; GENERAL INFORMATION:
; APPLICANT: Firestein, Stuart
; APPLICANT: Zhao, Haiqing
; TITLE OF INVENTION: Functional Expression of, and Assay for, Functional Cellular Re
; TITLE OF INVENTION: Vivo
; FILE REFERENCE: P01511US2 / 09805059
; CURRENT APPLICATION NUMBER: US/09/085,371
; CURRENT FILING DATE: 1998-05-19
; PRIOR APPLICATION NUMBER: US 08/891,243
; PRIOR FILING DATE: 1997-07-10
; PRIOR APPLICATION NUMBER: US 60/045,961
; PRIOR FILING DATE: 1997-05-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Rat
US-09-085-371-5

Query Match 17.3%; Score 84.4; DB 4; Length 900;
Best Local Similarity 48.4%; Pred. No. 7.4e-15;
Matches 235; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

```
Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||| || || ||| ||||  | |  || || || ||  |    | |   ||  | |  ||
Db     384 tgtggctatctgtcatccactccactaccccgtcattgtcagtagccggctatgtgtgca 443

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           | ||  |  |   || | |     |||||        |  ||||| | ||     |  |
Db     444 gatggcagctggatcctgggctggaggttttggtatctccatggttaaagttttccttat 503

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
           |  || ||  ||| |||  |    ||||  ||  |  | || | ||||||  | |  |
Db     504 ttctcgcctgtcttactgtggccccaacaccatcaaccacttttttctgtgatgtgtctcc 563

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           | || ||||  | |||||||  ||     ||  |     |  ||  | |   |||  |  | |
Db     564 attgctcaacctgtcatgcactgacatgtccacagcagagcttacagactttgtcctggc 623

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           |||||| |   |     |  |       |        || || |      |
Db     624 catttttattctgctgggaccgctctctgtcactggggcatcctacatggccatcacagg 683

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           || | |  ||   |     || |     ||| |    |||||||| || || || | | |
Db     684 tgctgtgatgcgcatcccctcagctgctggccgccataaagccttttcaacctgtgcctc 743

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||| ||   |  |||     | | ||| | | |    || |||| ||||   | |||
Db     744 ccacctcactgttgtgatcatcttctatgcagccagtattttcatctatgccaggcctaa 803

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           |   || ||||   |||    ||||  ||   ||| |  | ||| |  | ||  | ||
Db     804 ggcactctcagcttttgacaccaacaagctggtctctgtactctacgctgtcattgtacc 863

Qy     481 cctgct 486
           || |
Db     864 gttgtt 869
```

RESULT 8
US-08-748-506-7
; Sequence 7, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-7

Query Match 15.0%; Score 73.2; DB 3; Length 966;
Best Local Similarity 46.9%; Pred. No. 1.1e-11;
Matches 228; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          | | || ||||||   ||  | |  ||| ||      |||           |  | ||  |
Db    387 TATGGCTATATGTTCCCCACTCCACTATGCAACCCGAATGAGTCGTGGGGTATGTGCCTA 446

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           |||  || | | ||||  || || || | | |      |||||  ||    | |||  |
Db    447 TTTGGCAATTGTCTCATGGGTGATGGGATGCATAGTAGGTCTGGGACAGACCAATTTTAT 506

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||     | |  ||||| |          ||| |  | |||| ||||||  ||  ||
Db    507 TTTCTCCTTGAACTTCTGTGGACCCTGTGAGATAGACCACTTCTTCTGTGACCTTCCACC 566

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||  |   | || | ||   || | ||||    || |        || ||| |    |
Db    567 TCTCCTGGCACTTGCCTGTGGTGATACATCCCAAAATGAGGCTGCCATCTTTGTGGCAGC 626

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           |   |      | |     | ||  |    |||||    || |||    |  | ||| |
Db    627 AGTGCTCTGCATATTTAGTCCATTTTTACTGATCATTTCTTCCTATGTCAGAATTCTCGT 686

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           ||   | |||    |     ||     ||||  |    |||||  |||| || ||     |
Db    687 TGCAGTGCTGGTGATGCCTTCACCTGAGGGGCGCCATAAAGCTCTCTCTACCTGTTCATC 746

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           || || |||   ||| || | | ||| ||  |         ||  ||| || |  |
Db    747 TCACCTACTTGTAGTCACACTCTTCTATGGCTCAACATCTGCCACCTATTTGAGGTCCAA 806

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           ||| ||    |   ||     |||||| | |   | || || || ||     | |   |
Db    807 GTCTAGCCACTCACCAGGAGTGGACAAACTCTTGGCCCTCTTCTATACATCAGTGACATC 866

Qy    481 cctgct 486
           | ||||
Db    867 CATGCT 872


RESULT   9
US-08-748-506-9
; Sequence 9, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  984 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-9


  Query Match             14.7%;  Score 71.6;  DB 3;  Length 984;
  Best Local Similarity   46.7%;  Pred. No. 3.1e-11;
  Matches  227;  Conservative    0;  Mismatches  259;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          | | ||||| || || ||  | | |||||||   || |||  |   |   |  ||  |
Db    387 TCTGGCCATCTGCAAACCTCTACATTATCCAACCATCATGAGCCCAAGGATGTGCTTCCT 446

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           |  | | | | |   |    |  | || |||||            ||  ||     | ||
Db    447 TCTCGTTACTGTCTGTTTATTTTTGGGCTTCCTCTTCATGGCCAGTCCAGTTGTGATGCT 506

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           |       |  || |||  |    ||||| || | | | || | ||||||  ||  ||
Db    507 TTCCAAGACATTTTACTGTGGTCCAAACATTATTCCTCACTTTTTCTGTGATTTTGGACC 566

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||||   ||  | || ||    |   |     ||  |   |  | | ||||   |||
Db    567 ACTGGCAAATCTCTCCTGTTCAGAAACCAGGTCTATTGAGATGCTGTTTTTTACCCTTGC 626

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           ||   || | |   |  | | ||  | |    |||  |  | || |   |  | |  |
Db    627 TGTAATTGTGCTTTTTGCTTCCTTTCTTATAGCCATCTTTGCATACAGCAATATAGTAGT 686

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           ||  ||| |      || |  | ||    ||   | ||| || ||||| |||  | |
Db    687 CACCATAGTGAGACTCCCTTCAGCCAGGGAGCGACAGAGAGCTTTTTCCACCTGCTCCTC 746

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           |||||  ||    ||||  |    || || |        || || ||  ||   ||
Db    747 TCATCTCATTGTCCTCTCTCTAATGTATGGCAGCTGTGCATTTATATACCTGAAGCCAAA 806

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           |       |  |||    ||| || |  | ||||  |  ||| | || | |  ||
Db    807 GCAGAGAAGCAGAGTGGACACCAACAGAGAGGCTGCTCTTGTGAACATGGTTGTGACACC 866

Qy    481 cctgct 486
           ||| ||
Db    867 CCTTCT 872


RESULT  10
US-08-599-252-84/c
; Sequence 84, Application US/08599252
; Patent No. 5705343
;  GENERAL INFORMATION:
;    APPLICANT:  DRAYNA, DENNIS T.
;    APPLICANT:  FEDER, JOHN N.
;    APPLICANT:  GNIRKE, ANDREAS
;    APPLICANT:  KIMMEL, BRUCE E.
;    APPLICANT:  THOMAS, WINSTON J.
;    APPLICANT:  WOLFF, ROGER K.
;    TITLE OF INVENTION:  METHOD TO DIAGNOSE HEREDITARY
;    TITLE OF INVENTION:  HEMOCHROMATOSIS
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/599,252
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  9053-0001.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  84:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1320 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-599-252-84


  Query Match             10.6%;  Score 51.4;  DB 1;  Length 1320;
  Best Local Similarity   61.7%;  Pred. No. 1.7e-05;
  Matches   82;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy    349 cacatgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgta 408
          ||| ||| || | |||||  ||  ||||||  | | ||| || ||  | || | ||||||
Db   1320 CACTTGCAGCTCGCATCTAATTGTTGTCTCTCTCTTCTATGGTACAATCATATACATGTA 1261

Qy    409 tgtgcgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttacac 468
            | |  || |||  |   |   | ||||| | |||  |   | | || || |||||
Db   1260 CCTCCAGCCAGCAAATACTTATTCCCAGGACCAGGGCAAGTTTCTTACCCTTTTCTACAC 1201

Qy    469 gattataattccc 481
            ||| | | ||||
Db   1200 AATTGTCACTCCC 1188


RESULT  11
US-08-436-074-57/c
; Sequence 57, Application US/08436074
; Patent No. 5753438
;  GENERAL INFORMATION:
;    APPLICANT:  DRAYNA, DENNIS T.
;    APPLICANT:  FEDER, JOHN N.
;    APPLICANT:  GNIRKE, ANDREAS
;    APPLICANT:  KIMMEL, BRUCE E.
;    APPLICANT:  THOMAS, WINSTON J.
;    APPLICANT:  WOLFF, ROGER K.
;    TITLE OF INVENTION:  METHOD TO DIAGNOSE HEREDITARY
;    TITLE OF INVENTION:  HEMOCHROMATOSIS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,074
;      FILING DATE:  08-MAY-1995
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  9053-0001.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1320 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-436-074-57


  Query Match             10.6%;  Score 51.4;  DB 1;  Length 1320;
  Best Local Similarity   61.7%;  Pred. No. 1.7e-05;
  Matches   82;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy    349 cacatgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgta 408
          ||| ||| || | |||||  ||  ||||||  | | ||| || ||  | || | ||||||
Db   1320 CACTTGCAGCTCGCATCTAATTGTTGTCTCTCTCTTCTATGGTACAATCATATACATGTA 1261

Qy    409 tgtgcgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttacac 468
            | |  || |||  |   |   | ||||| | |||  |   | | || || |||||
Db   1260 CCTCCAGCCAGCAAATACTTATTCCCAGGACCAGGGCAAGTTTCTTACCCTTTTCTACAC 1201

Qy    469 gattataattccc 481
            ||| | | ||||
Db   1200 AATTGTCACTCCC 1188


RESULT  12
PCT-US96-06352-84/c
; Sequence 84, Application PC/TUS9606352
;  GENERAL INFORMATION:
;    APPLICANT:  DRAYNA, DENNIS T.
;    APPLICANT:  FEDER, JOHN N.
;    APPLICANT:  GNIRKE, ANDREAS
;    APPLICANT:  KIMMEL, BRUCE E.
;    APPLICANT:  THOMAS, WINSTON J.
;    APPLICANT:  WOLFF, ROGER K.
;    TITLE OF INVENTION:  METHOD TO DIAGNOSE HEREDITARY
;    TITLE OF INVENTION:  HEMOCHROMATOSIS
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/06352
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,252
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  9053-0001.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  84:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1320 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US96-06352-84


  Query Match            10.6%;  Score 51.4;  DB 5;  Length 1320;
  Best Local Similarity   61.7%;  Pred. No. 1.7e-05;
  Matches   82;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy     349 cacatgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgta 408
           ||| ||| || | |||||  ||  ||||||  | | ||| || ||  | || | ||||||
Db    1320 CACTTGCAGCTCGCATCTAATTGTTGTCTCTCTCTTCTATGGTACAATCATATACATGTA 1261

Qy     409 tgtgcgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttacac 468
             | |  || ||| |   |   |  | |||| | |||  |   | | || || |||||
Db    1260 CCTCCAGCCAGCAAATACTTATTCCCAGGACCAGGGCAAGTTTCTTACCCTTTTCTACAC 1201

Qy     469 gattataattccc 481
           ||| | | ||||
Db    1200 AATTGTCACTCCC 1188


RESULT  13
PCT-US96-06583-84/c
; Sequence 84, Application PC/TUS9606583
;  GENERAL INFORMATION:
;    APPLICANT:  DRAYNA, DENNIS T.
;    APPLICANT:  FEDER, JOHN N.
;    APPLICANT:  GNIRKE, ANDREAS
;    APPLICANT:  KIMMEL, BRUCE E.
;    APPLICANT:  THOMAS, WINSTON J.
;    APPLICANT:  WOLFF, ROGER K.
;    TITLE OF INVENTION:  METHOD TO DIAGNOSE HEREDITARY
;    TITLE OF INVENTION:  HEMOCHROMATOSIS
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/06583
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,252
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  9053-0001.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  84:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1320 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US96-06583-84


  Query Match            10.6%;  Score 51.4;  DB 5;  Length 1320;
  Best Local Similarity   61.7%;  Pred. No. 1.7e-05;
  Matches   82;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy     349 cacatgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgta 408
           ||| ||| || | |||||  ||  ||||||  | | ||| || ||  | || | ||||||
Db    1320 CACTTGCAGCTCGCATCTAATTGTTGTCTCTCTCTTCTATGGTACAATCATATACATGTA 1261

Qy     409 tgtgcgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttacac 468
             | |  || ||| |   |   |  | |||| | |||  |   | | || || |||||
Db    1260 CCTCCAGCCAGCAAATACTTATTCCCAGGACCAGGGCAAGTTTCTTACCCTTTTCTACAC 1201

Qy     469 gattataattccc 481
           ||| | | ||||
Db    1200 AATTGTCACTCCC 1188


RESULT  14
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             9.3%;  Score 45.4;  DB 1;  Length 7218;
  Best Local Similarity   6.0%;  Pred. No. 0.0015;
  Matches   25;  Conservative  212;  Mismatches  178;  Indels    0;  Gaps    0;

Qy      52 cagcgctcagttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgt 111
           |||  |   |  |||    ||::: : : :  ::  :::::: : :::::::  ::: : :
Db    1046 CAGGTCGAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1105
```

```
Qy    112 gagtttactattgcgactaactttctgcaggtttaacataatacattatttctactgtga 171
          ::: :: :: :  ::  :::::: :  :::  : :  : : :: ::::: :: :
Db   1106 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1165

Qy    172 aattttacaactgttcaaaatttcatgcaatggtccatctattaacgcactaataatatt 231
           :::: :  :: :::    :::: : :  :  ::: ::: ::  : : ::  :  : ::
Db   1166 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1225

Qy    232 tatttttggtgcttttatacaaatacccactttaatgactatcataatctcttatactcg 291
          : :::::  : ::::: : :   : ::: :::  :  :: :: :  :::::: : :::
Db   1226 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1285

Qy    292 tgtgctctttgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccac 351
          : : ::::::  : ::::             :::      :    :    :::::::: :
Db   1286 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1345

Qy    352 atgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgt 411
           : :  : ::: ::: :::::: :::: :: : :: :    ::::  ::::: : : : :
Db   1346 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1405

Qy    412 gcgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttac 466
          : :::: : :::  :::  ::      ::    |||| | |  | ||| ||| ||
Db   1406 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCAAATTCTTCTATCTCTTTAAC 1460
```

```
RESULT  15
US-08-991-677-11/c
; Sequence 11, Application US/08991677A
; Patent No. 6252135
; GENERAL INFORMATION:
;  APPLICANT: Chiang, Vincent L
;  APPLICANT: Carraway, Daniel T
;  APPLICANT: Smeltzer, Richard H
;  TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
;  FILE REFERENCE: 50617
;  CURRENT APPLICATION NUMBER: US/08/991,677A
;  CURRENT FILING DATE: 1997-12-16
;  EARLIER APPLICATION NUMBER: US 60/033,381
;  EARLIER FILING DATE: 1996-12-16
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 2251
;   TYPE: DNA
;   ORGANISM: Pinus taeda
US-08-991-677-11
```

```
  Query Match             7.9%;  Score 38.4;  DB 4;  Length 2251;
  Best Local Similarity   49.5%;  Pred. No. 0.094;
  Matches   99;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;

Qy    118 actattgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttt 177
          | |||| || ||  |  || ||    |    |||| | |  |||  |  | |  || |||
Db    889 AATATTTGGAATATTTGACTTCAAAAGTTTATTAATTTAAAATTAATTTTTTCTAAATTT 830

Qy    178 acaactgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttt 237
            |  |    | ||  | |         ||  | ||     | ||||| | |||| ||
Db    829 TAACTTAAATATAAGATAAAATTTACCAACAAATTTTTTAAAAATAATAGTGTTTAATTA 770

Qy    238 tggtgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgct 297
          |  |   |||  | |  |   | ||  | | |  |  | ||| | |||| |   | |  |
Db    769 TTTTAGATTTTGAAATTTTTACTCTAAATTTAACAAAACAATTTTTTATCCAATTATAAT 710

Qy    298 ctttgatattctgaaaaaaa 317
           |||| |||||  || ||||
Db    709 TTTTGTTATTCAAAATAAAA 690
```

```
Search completed: June 27, 2002, 11:17:06
Job time: 3519 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 10:18:27 ; Search time 1614.01 Seconds
(without alignments)
4072.476 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	228.2	46.9	757	12	BH278852	BH278852 CH230-108
	2	177.4	36.4	623	12	AZ641407	AZ641407 1M0503K20
c	3	171.2	35.2	496	12	AZ366926	AZ366926 1M0116D24
c	4	165.2	33.9	681	12	AG136576	AG136576 Pan trogl
	5	160.8	33.0	570	12	AZ090010	AZ090010 RPCI-23-4
	6	150.8	31.0	404	12	AZ880560	AZ880560 RPCI-23-1
c	7	141.8	29.1	629	12	AZ019257	AZ019257 RPCI-23-2
c	8	138.2	28.4	550	12	AQ984439	AQ984439 RPCI-23-3
c	9	129.8	26.7	556	12	AZ714202	AZ714202 RPCI-24-8
c	10	129.2	26.5	649	12	BH342817	BH342817 CH230-64F
c	11	126.8	26.0	763	12	BH109232	BH109232 RPCI-24-3
c	12	126	25.9	611	12	AZ103967	AZ103967 RPCI-23-3
c	13	124.4	25.5	507	12	AZ560017	AZ560017 RPCI-23-2
c	14	124.4	25.5	628	12	AZ086388	AZ086388 RPCI-23-2
c	15	118.2	24.3	740	12	AZ607393	AZ607393 1M0429N15
	16	115	23.6	508	12	AZ373636	AZ373636 1M0126C20
	17	112.4	23.1	682	12	AZ334227	AZ334227 1M0063H01
	18	111.6	22.9	2021	11	BC016940	BC016940 Homo sapi
	19	110.8	22.8	610	12	AZ392378	AZ392378 1M0154P24
	20	110	22.6	642	12	AZ969227	AZ969227 2M0241J24
c	21	109.8	22.5	656	12	BH078380	BH078380 RPCI-24-3
	22	106.2	21.8	580	12	AZ418115	AZ418115 1M0194N03
c	23	106	21.8	676	12	AG069165	AG069165 Pan trogl
	24	105.2	21.6	646	12	AZ638594	AZ638594 1M0498M06
c	25	102.6	21.1	675	12	AZ430777	AZ430777 1M0215H24
c	26	101	20.7	587	12	AZ461407	AZ461407 1M0267F18
c	27	100.2	20.6	438	12	AQ984812	AQ984812 RPCI-23-2
c	28	99.4	20.4	578	12	AZ252633	AZ252633 RPCI-23-7
c	29	99.2	20.4	652	12	AZ972672	AZ972672 2M0246D10
	30	98.2	20.2	732	10	BG068751	BG068751 H3068H10-
c	31	97.2	20.0	686	12	AZ086625	AZ086625 RPCI-23-2
	32	95.6	19.6	513	12	AZ062239	AZ062239 RPCI-23-4
	33	95.4	19.6	726	12	AZ255734	AZ255734 RPCI-23-1
	34	95.2	19.5	678	12	BH067899	BH067899 RPCI-24-3
c	35	95	19.5	515	12	AZ779071	AZ779071 2M0014L24
	36	94	19.3	580	12	AZ593814	AZ593814 1M0405K03
c	37	93.8	19.3	711	12	BH291047	BH291047 CH230-103
	38	93.6	19.2	647	12	AZ396019	AZ396019 1M0160I23
c	39	93.4	19.2	294	12	AQ027654	AQ027654 CIT-HSP-2
c	40	93.4	19.2	543	12	AZ742029	AZ742029 RPCI-24-9
	41	93.2	19.1	236	12	BH405788	BH405788 RPCI-23-2
	42	92.8	19.1	1394	11	AK017036	AK017036 Mus muscu
	43	92.8	19.1	3063	11	AK016560	AK016560 Mus muscu
c	44	92.6	19.0	639	12	BH078443	BH078443 RPCI-24-3
c	45	92.4	19.0	698	12	BH069789	BH069789 RPCI-24-3

ALIGNMENTS

RESULT 1
BH278852/c
LOCUS BH278852 757 bp DNA linear GSS 30-NOV-2001
DEFINITION CH230-108O16.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone CH230-108O16, DNA sequence.
ACCESSION BH278852
VERSION BH278852.1 GI:17191254
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 757)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: CH230-108O16.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or ering_information.htm). BAC end page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 108 row: O column: 16
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .757
/organism="Rattus norvegicus"
/strain="BN/SsNHsd/MCW"

```
                     /db_xref="taxon:10116"
                     /clone="CH230-108O16"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      292 a    110 c    137 g    218 t
ORIGIN


  Query Match           46.9%;  Score 228.2;  DB 12;  Length 757;
  Best Local Similarity   67.2%;  Pred. No. 1.3e-43;
  Matches  323;  Conservative    0;  Mismatches  158;  Indels    0;  Gaps    0;

Qy      6 ccatatgtaatccctttgctttatccagtgatgatgtccaacaaactcagcgctcagttgc 65
          | ||||| || || || ||||| ||||| || ||| | || |  ||| | |  |   | |
Db    757 CTATATGCAAACCTTTACTTTACCCAGTCATCATGACAAATAGGCTCTGTGTACGTCTTC 698

Qy     66 taagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactattgc 125
          |||  || || | | ||  ||| ||  | ||| || |  ||||||     ||  ||||
Db    697 TAACAATGTCCTTTATAGGTGGATTTATTCATACTTTCATTCATGCAGTCTTTTTATTCA 638

Qy    126 gactaactttctgcaggtttaacataatacattatttctactgtgaaatttttacaactgt 185
          ||||||| |||||    | ||||||||||| |  | || || ||||| ||| | | |||||
Db    637 GACTAACATTCTGTGATTCTAACATAATATACCACTTTTATTGTGATATTATGCCACTGT 578

Qy    186 tcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggtgctt 245
          | || ||||| || | ||  || ||| | ||    || || || ||||||||   ||  |
Db    577 TAAAGATTTCCTGTAGTGACCCTTCTCTCAATTACCTGATGATTTTTATTTTCTCTGGCT 518

Qy    246 ttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctttgata 305
            ||| ||  || | |  |||| ||  | ||||||||||| | ||| || ||    |
Db    517 CAATTCAAGTATTCAGTATTTTGACCATTCTTATCTCTTATACGCTTGTTCTGTTCTCAA 458

Qy    306 ttctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcccatc 365
          | || ||  | || |||   || |||| ||  ||||| || ||||| ||||| |||||||
Db    457 TCCTAAAGCAGAAATCTATCAAAGGCATAAAGAAAGCTTTATCCACCTGCGGAGCCCATC 398

Qy    366 tgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcatctg 425
          | || ||||| || || ||||| ||  |||| ||||||||||| ||||||||||||||
Db    397 TCCTCTCTGTATCTTTATACTATGGCTCTCTTCTCTTCATGTATTTGCGTCCTGCATCTC 338

Qy    426 gcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcccctgc 485
            |  || || ||||| |   || | ||| | |||||||  || |||||||  | |
Db    337 CGCATGTAGATGATCAAGATATGATGGACTCTGTATTTTACACTGTTTTAATTCCTGTAC 278

Qy    486 t 486
          |
Db    277 T 277


RESULT   2
AZ641407
LOCUS       AZ641407                623 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0503K20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0503K20 R, DNA sequence.
ACCESSION   AZ641407
VERSION     AZ641407.1  GI:11765357
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 623)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0503  row: K  column: 20
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 623.
FEATURES             Location/Qualifiers
     source          1. .623
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0503K20"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      161 a    142 c    117 g    203 t
ORIGIN


  Query Match           36.4%;  Score 177.4;  DB 12;  Length 623;
  Best Local Similarity   72.7%;  Pred. No. 1.1e-31;
  Matches  229;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy    173 attttacaactgttcaaaatttcatgcaatggtccatctattaacgcactaataatattt 232
          || || ||  | |||||||||||| |||| || |||  | || ||||  ||  |  | ||
Db      1 ATCTTGCAGTTATTCAAAATTTCCTGCACTGATCCTACAATGAACGTCCTTCTGGTTTTA 60

Qy    233 atttttggtgcttttatacaaatacccactttaatgactatcataatctcttatactcgt 292
          || |||   ||||| ||||||      ||||| |||||||| ||  |||| ||      |
Db     61 ATATTTTCAGCTTTGATACAAGGCTTGACTTTTATGACTATTATCGTCTCCTACTTCTCT 120

Qy    293 gtgctctttgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccaca 352
          || ||  |||  || |||||||| ||||||||| | ||||||| ||||||||||||||||
Db    121 GTCCTTCTTGCCATCCTGAAAACAAAGTCTGAGAGGGGCAGACGCAAAGCCTTCTCCACC 180

Qy    353 tgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtg 412
          ||| | || || ||||| || || ||  ||| ||| || || ||  || ||||||||||
Db    181 TGCAGTGCTCACCTGCTCTCCGTGTCTCTGTTCTATGGCACACTTTTCCTCATGTATGTG 240

Qy    413 cgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttacacgatt 472
          ||||| |  ||||| | || |||||  | |||| | ||||||| | ||||||| ||
Db    241 CGTCCCGGGTCTGGATCAGGTGAAGATAAGGACAGAATGTATTCTTTATTTTACACAATA 300

Qy    473 ataattcccctgcta 487
          ||||||||| | |||
Db    301 ATAATTCCCTTTCTA 315
```

```
RESULT   3
AZ366926/c
LOCUS       AZ366926                 496 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0116D24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0116D24 F, DNA sequence.
ACCESSION   AZ366926
VERSION     AZ366926.1  GI:10480626
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0116  row: D  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 496.
FEATURES             Location/Qualifiers
     source          1. .496
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0116D24"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      171 a     80 c    110 g    135 t
ORIGIN


  Query Match            35.2%;  Score 171.2;  DB 12;  Length 496;
  Best Local Similarity  71.7%;  Pred. No. 3.2e-30;
  Matches  238;  Conservative    0;  Mismatches   93;  Indels    1;  Gaps    1;

Qy    157 ttatttctactgtgaaattttacaactgttcaaaatttcatgcaatg-gtccatctatta 215
          | |||| |||||||| ||  | | ||| |  | |||||| ||||  |   ||||||||||
Db    496 TCATTTTTACTGTGAGATCCTGCCACTCTATACAATTTCTTGCACCGACCCCATCTATTA 437

Qy    216 acgcactaataatatttatttttggtgcttttatacaaatacccactttaatgactatca 275
          | ||| |  |  | || |||||  |||| | ||||||      || || |||| |||
Db    436 ATGCATTTGTGGTCTTCATTTTTTCTGCTGTGATACAAGCTGTTACCTTCATGAGTATTG 377

Qy    276 taatctcttatactcgtgtgctctttgatattctgaaaaaaaagtctgaaaagggcagaa 335
          | |||| ||| | |  || ||||||   || ||||||| |||||||| | ||||||||
Db    376 CAGTCTCCTATGCCCACGTCCTCTTTTCCATCCTGAAAACAAAGTCTGAGAGGGGCAGAA 317

Qy    336 gcaaagccttctccacatgcggcgcccatctgctttctgtctcattgtactacggaactc 395
          | |||||||||||| | ||| | ||||| || || ||||||||  ||| ||| ||||||||
Db    316 GAAAAGCCTTCTCCCCCTGCAGTGCCCACCTACTCTCTGTCTCTCTGTTCTATGGAACTC 257

Qy    396 tgatcttcatgtatgtgcgtcctgcatctggcttagctgaagaccaagacaaagtgtatt 455
          |  |||| ||||| ||| | |||| |||||   |  | || |  |  | ||  ||||||
Db    256 TCTTCTTTATGTACGTGAGCCCTGGGTCTGGACCAAGTAAATATAAGAATAAGATGTATT 197

Qy    456 ctctgttttacacgattataattcccctgcta 487
          |||||||||||||| ||| | ||||| || |||
Db    196 CTCTGTTTTACACCATTGTGATTCCTCTACTA 165


RESULT   4
AG136576/c
LOCUS       AG136576                 681 bp    DNA     linear   GSS 04-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-150D04.R, genomic survey sequence.
ACCESSION   AG136576
VERSION     AG136576.1  GI:16666254
KEYWORDS    GSS; GSS (genome survey sequence).
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-150D04.R.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 681)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="PTB-150D04.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      247 a    117 c    143 g    173 t      1 others
ORIGIN


  Query Match            33.9%;  Score 165.2;  DB 12;  Length 681;
  Best Local Similarity  64.8%;  Pred. No. 8.2e-29;
  Matches  245;  Conservative    0;  Mismatches  133;  Indels    0;  Gaps    0;
```

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||| || || || |||||||||||| || ||| ||||   |||  ||
Db    426 TGTAGCCATATGCAAACCTTTACTTTATCCAGTCATTATGACCAATGGACTGTGCATCTG 367

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          | |  |||   | |||| | ||  |||  | || ||| || |  ||||||   ||||  |
Db    366 GCTATTAATCTTGTCATTTCTAGGTGGCCTTCTTCATGCTTTAATTCATGAAGGTTTTTT 307

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          |    || |||| ||||| |  |  |||| ||||||| | || ||||||| ||| | |
Db    306 AAACAGATTAACCTTCTGTAATTCCAACACAATACATCACTTTTACTGTGACATTATCCC 247

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          | |||| |||||||  || | || | | |||| ||||   | |||  | ||||||||
Db    246 ATTGTTAAAAATTTTCTGTACTGATTCTTCTACTAACTTTCCAATGGTTTTTATTTTCTC 187

Qy    241 tgctttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
            ||  || ||| |   |||  |   |||| |  |  | ||||||||   ||| ||||
Db    186 ATGTTCAATTCAAGTTTTCACCATTGGGACTGTTCTTGTATCTTATACATTTGTCCTCTC 127

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |   ||  ||||||| ||||||  || | || ||| |||| |||  |||| || || ||
Db    126 TACAATCTTGAAAAAGAAGTCTGTCAAAGACATAAGAAAAGCGTTCACCACCTGTGGAGC 67

Qy    361 ccatctgctttctgtctc 378
          |    || | ||  | |
Db     66 TCTCGAGCATGCTAGCGC 49


RESULT  5
AZ090010
LOCUS       AZ090010                570 bp    DNA     linear   GSS 08-MAY-2000
DEFINITION  RPCI-23-470I16.TV RPCI-23 Mus musculus genomic clone RPCI-23-470I16
            , DNA sequence.
ACCESSION   AZ090010
VERSION     AZ090010.1  GI:7732053
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-470I16.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 470  row: I  column: 16
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-470I16"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT      155 a    108 c     88 g    219 t
ORIGIN


  Query Match           33.0%;  Score 160.8;  DB 12;  Length 570;
  Best Local Similarity   64.5%;  Pred. No. 8.9e-28;
  Matches  240;  Conservative    0;  Mismatches  132;  Indels    0;  Gaps    0;

Qy    116 ttactattgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatt 175
          || ||||| | ||||| ||||| |  |  || ||| ||||| | || | |||||| ||
Db      2 TTTCTATTACTACTAAACTTCTGTAATTCCAATATAGTACATCACTTTTTCTGTGACATC 61

Qy    176 ttacaactgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatt 235
          | | | || | || ||||| || | ||    |  || ||||    | | | |||||| ||
Db     62 GTTCCATTGCTAAAGATTTCCTGTACTGACACTACTCTTAATTTTCAACTGATATTTGTT 121

Qy    236 tttggtgcttttatacaaatacccacttaatgactatcataatctcttatactcgtgtg 295
          || | ||   | || ||| | | |||  |  ||| | |  | || ||||||| | ||
Db    122 TTAGCTGGAATAATTCAAGTCCTCACCGTTGTGATTGTTCTTGTGTCCTATACTCTTTTG 181

Qy    296 ctctttgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgc 355
          || ||    ||| || |||  |||||||   |  |||  |  || || ||||| || ||
Db    182 CTATTGACGATTTTGCAAAGGAAGTCTGTCCAAAGCATGAAGAAGGCTTTCTCTACCTGT 241

Qy    356 ggcgcccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgt 415
          || |||||||  | ||||| ||  ||||||| ||  ||||  || |||||||| |
Db    242 GGTGCCCATCTCTTATCTGTGTCTCTGTACTATGGGCCTCTTCTCATCATGTATATTTTC 301

Qy    416 cctgcatctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattata 475
          || |  ||      ||| || |  ||||| |   |  | ||||||||||||||| | |||
Db    302 CCCGTGTCCCAAGAAGCAGATGGTCAAGATATCATAGACTCTCTGTTTTACACGGTAATA 361

Qy    476 attccccctgcta 487
          |||||  || ||
Db    362 ATTCCTGTGTTA 373


RESULT  6
AZ880560
LOCUS       AZ880560                404 bp    DNA     linear   GSS 05-MAR-2001
DEFINITION  RPCI-23-185I4.TJ RPCI-23 Mus musculus genomic clone RPCI-23-185I4,
            DNA sequence.
ACCESSION   AZ880560
VERSION     AZ880560.1  GI:13199505
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 404)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-185I4.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
```

```
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 185  row: I  column: 4
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .404
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-23-185I4"
                      /clone_lib="RPCI-23"
                      /sex="Female"
                      /lab_host="DH10B"
                      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                      brain genomic DNA was isolated and  partially digested
                      with a combination of EcoRI and EcoRI Methylase.   Size
                      selected DNA was cloned into the pBACe3.6 vector at the
                      EcoRI  sites.  The ligation products were transformed into
                      DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT       99 a     90 c     67 g    148 t
ORIGIN


  Query Match             31.0%;  Score 150.8;  DB 12;  Length 404;
  Best Local Similarity   63.9%;  Pred. No. 2e-25;
  Matches  244;  Conservative    0;  Mismatches  137;  Indels    1;  Gaps    1;

Qy    106 tcatgtgagtttactattgcgactaactttctgcaggtttaacataatacattatttcta 165
          ||||| | | || ||||| | |||||  ||||| |  |  || ||| ||||| | || |
Db      1 TCATGAGGGATTTCTATTACTACTAAACTTCTGTAATTCCAATATAGTACATCACTTTTT 60

Qy    166 ctgtgaaattttacaactgttcaaaatttcatgcaatggtccatctattaacgcactaat 225
          |||||| ||  | | | || | || ||||| || | ||   |  || ||||    | | |
Db     61 CTGTGACATCGTTCCATTGCTAAAGATTTCCTGTACTGACACTACTCTTAATTTTCAACT 120

Qy    226 aatatttatttttggtgcttttatacaaatacccactttaatgactatcataatctctta 285
           |||||| |||| | ||   | || ||| | | |||  |  ||| | |  |  | || |
Db    121 GATATTTGTTTTAGCTGGAATAATTCAAGTCCTCACCGTTGTGATTGTTCTTGTGTC-TC 179

Qy    286 tactcgtgtgctctttgatattctgaaaaaaaagtctgaaaagggcagaagcaaagcctt 345
          ||| |  ||||| ||    ||| || |||  |||||||   |  ||| |  || || ||
Db    180 TACACTAGTGCTATTGACGATTTTGCAAAGGAAGTCTGTCCAAAGCATGAAGAAGGCTTT 239

Qy    346 ctccacatgcggcgcccatctgctttctgtctcattgtactacggaactctgatcttcat 405
          ||| || || || |||||||| | ||||| ||  |||||||  ||  ||||  || ||||
Db    240 CTCTACCTGTGGTGCCCATCTCTTATCTGTGTCTCTGTACTATGGGCCTCTTCTCATCAT 299

Qy    406 gtatgtgcgtcctgcatctggcttagctgaagaccaagacaaagtgtattctctgtttta 465
          |||| |    || |  ||       ||| || |  ||||| |   |  | |||||||| ||
Db    300 GTATATTTTCCCCGTGTCCCAAGAAGCAGATGGTCAAGATATCATAGACTCTCTGTTCTA 359

Qy    466 cacgattataattcccctgcta 487
          |||| | ||||||||  || ||
Db    360 CACGGTCATAATTCCTGTGTTA 381


RESULT   7
AZ019257/c
LOCUS       AZ019257                629 bp    DNA     linear   GSS 25-FEB-2000
DEFINITION  RPCI-23-266H21.TV RPCI-23 Mus musculus genomic clone RPCI-23-266H21
            , DNA sequence.
ACCESSION   AZ019257
VERSION     AZ019257.1  GI:7094641
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-266H21.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 266  row: H  column: 21
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .629
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-23-266H21"
                      /clone_lib="RPCI-23"
                      /sex="Female"
                      /lab_host="DH10B"
                      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                      brain genomic DNA was isolated and  partially digested
                      with a combination of EcoRI and EcoRI Methylase.   Size
                      selected DNA was cloned into the pBACe3.6 vector at the
                      EcoRI  sites.  The ligation products were transformed into
                      DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      218 a    124 c    135 g    151 t      1 others
ORIGIN


  Query Match             29.1%;  Score 141.8;  DB 12;  Length 629;
  Best Local Similarity   55.8%;  Pred. No. 2.6e-23;
  Matches  269;  Conservative    0;  Mismatches  213;  Indels    0;  Gaps    0;

Qy      5 gccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcagttg 64
          ||  | || || ||||| || ||    || || |||||||| |   || | | |||||
Db    567 GCTGTCTGCAAACCCTTACTGTACAGTGTCATTATGTCCAAGAGGGTCTGTGTTCAGTNT 508

Qy     65 ctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactattg 124
          | |      | ||| |  | | |    |  |          ||||   |  ||    ||
Db    507 GTCATAGGACCTTATCTTTTAGCTGCTATAAACATCACAACTCATACAACATTGACCTTT 448

Qy    125 cgactaactttctgcaggtttaacataatacattatttctactgtgaaatttacaactg 184
          |  || | ||||  | |  || |  ||  |  |||||| ||||||  |||  | ||||
Db    447 TGCTTACCGTTCTGTGGTTCAAATACCATCAACCATTTCTTCTGTGATGTTTCCCCACTG 388

Qy    185 ttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggtgct 244
          |   | |  ||||   ||   |||  ||||||    || || | ||| |  | | ||
Db    387 CTTTCATTAGCATGTGCTGACACATGGATTAACAAGGTAGTACTTTTTGTCCTGGCTGGA 328

Qy    245 tttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctttgat 304
          | |||    | |||| |   |||  ||||||  | || |||    |  | ||  | | |
Db    327 GTGATAGGTGTTCCCAGTGGCTTGATAATCATAGTTTCCTATGTGTGCATCCTGGTGGCT 268

Qy    305 attctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcccat 364
          ||  |||| |  ||| |||   | || |||   ||||||||||||||| ||     | ||
Db    267 ATCTTGAAGATCAAGACTGCTGATGGGAGACAGAAAGCCTTCTCCACGTGTTCTTCTCAC 208

Qy    365 ctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcatct 424
          |||    ||||  || |    || || |||||  ||||||| ||||| || |||    |
Db    207 CTGGCAGCTGTAACAATTCTTTATGGCACTCTTTTCTTCATATATGTTCGGCCTAATGCA 148

Qy    425 ggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcccctg 484
```

```
Db    147 AGTTCTTCCCTAGATATTAATAAAGTGATCTCTTTATTTTATACGGTGGTGATCCCCATG 88

Qy    485 ct 486
Db     87 TT 86


RESULT  8
AQ984439/c
LOCUS       AQ984439                550 bp    DNA     linear   GSS 30-JAN-2000
DEFINITION  RPCI-23-322E11.TJ RPCI-23 Mus musculus genomic clone RPCI-23-322E11
            , DNA sequence.
ACCESSION   AQ984439
VERSION     AQ984439.1  GI:6817644
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-322E11.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 322  row: E  column: 11
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-322E11"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      189 a    110 c    115 g    136 t
ORIGIN


  Query Match           28.4%;  Score 138.2;  DB 12;  Length 550;
  Best Local Similarity   56.5%;  Pred. No. 1.8e-22;
  Matches  257;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy     32 gtgatgatgtccaacaaactcagcgctcagttgctaagtatttcatatgtaattggtttc 91
Db    548 GTCATTATGTCCAAGAGGGTCTGTGTTCAGTTTGTCATAGGACCTTATCTTTTAGCTGCT 489

Qy     92 ctgcatcctctggttcatgtgagtttactattgcgactaactttctgcaggtttaacata 151

Db    488 ATAAACATCCCAACTCATACAACATTGACCTTTTGCTTACCGTTCTGTGGTTCAAATACC 429

Qy    152 atacattatttctactgtgaaattttacaactgttcaaaatttcatgcaatggtccatct 211
Db    428 ATCAACCATTTCTTCTGTGATGTTTCCCCACTGCTTTCATTAGCATGTGCTGACACATGG 369

Qy    212 attaacgcactaataatatttatttttggtgcttttatacaaatacccactttaatgact 271
Db    368 ATTAACAAGGTAGTACTTTTTGTCCTGGCTGGAGTGATAGGTGTTCCCAGTGGCTTGATA 309

Qy    272 atcataatctcttatactcgtgtgctctttgatattctgaaaaaaaagtctgaaaagggc 331
Db    308 ATCATAGTTTCCTATGTGTGCATCCTGGTGGCTATCTTGAAGATCAAGACTGCTGATGGG 249

Qy    332 agaagcaaagccttctccacatgcggcgcccatctgctttctgtctcattgtactacgga 391
Db    248 AGACAGAAAGCCTTCTCCACGTGTTCTTCTCACCTGGCAGCTGTCTCAATTCTTTATGGC 189

Qy    392 actctgatcttcatgtatgtgcgtcctgcatctggcttagctgaagaccaagacaaagtg 451
Db    188 ACTCTTTTCTTCATATATGTTCGGCCTAATGCAAGTTCTTCCCTAGATATTAATAAAGTG 129

Qy    452 tattctctgttttacacgattataattcccctgct 486
Db    128 ATCTCTTTATTTTATACGGTGGTGATCCCCATGTT 94


RESULT  9
AZ714202/c
LOCUS       AZ714202                556 bp    DNA     linear   GSS 24-JAN-2001
DEFINITION  RPCI-24-86K20.TJ RPCI-24 Mus musculus genomic clone RPCI-24-86K20,
            DNA sequence.
ACCESSION   AZ714202
VERSION     AZ714202.1  GI:12449680
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 86  row: K  column: 20
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-86K20"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
```

```
                    DNA."
BASE COUNT      184 a    103 c    136 g    133 t
ORIGIN
```

```
  Query Match             26.7%;  Score 129.8;  DB 12;  Length 556;
  Best Local Similarity   54.5%;  Pred. No. 1.7e-20;
  Matches  260;  Conservative    0;  Mismatches  217;  Indels    0;  Gaps    0;

Qy      11 tgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcagttgctaagt 70
           ||||| |||||||| ||     |  |||||||| |  || || || |   | || ||  |
Db     535 TGTAACCCCTTGCTATACAATGCTATGATGTCCTATCAAGTCTGCACCTGGATGATATTT 476

Qy      71 atttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactattgcgacta 130
             |||   || |||||  |   | |    | |||       |   |  |  || |
Db     475 GGGGTGTATAGCATGGGTTTTATTGGTGCCACAGCTCACACAGTGTGCATGCTAAGAGTC 416

Qy     131 actttctgcaggtttaacataatacattatttctactgtgaaattttacaactgttcaaa 190
           |||||| | | || |  ||||| || ||| || ||||||  |||| |   | ||  |
Db     415 CATTTCTGTAAGGTTGATGTAATAAATCATTACTTCTGTGATCTTTTCCCGTTATTGGAG 356

Qy     191 atttcatgcaatggtccatctattaacgcactaataatatttatttttggtgcttttata 250
           |||| |||     | | | ||| || |   || |  |  |   |||| |||||||
Db     355 CTTTCTTGCTCCCCTACTTTTATCAATGAGGTAGTTGTTCTATGTTTTAGTGCTTTCAAT 296

Qy     251 caaatacccactttaatgactatcataatctcttatactcgtgtgctctttgatattctg 310
             |   | |   | ||| |||| | | |||||| |       |  |       || ||
Db     295 ATCCTTTTCCCAACACTGAGTATCCTGAGCTCTTACATCTTCATCATTGCCAGCATCCTC 236

Qy     311 aaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcccatctgctt 370
             | || ||     | |||||  ||||||||||  ||| ||| || | ||  |
Db     235 CGTATTAAATCCACTGAAGGCAGGTCCAAAGCCTTCAGCACCTGCAGCTCACACATAACA 176

Qy     371 tctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcatctggctta 430
           |||   |  ||| ||  ||  ||      || |||||  | |  ||  |||| | |
Db     175 ACTGATGCTGTGTTCTTTGGGTCTGCTGCATTTATGTACCTACAGCCATCATCAGTCAGC 116

Qy     431 gctgaagaccaagacaaagtgtattctctgttttacacgattataattcccctgcta 487
           |    |||||||  ||||| |  ||| | ||||| || ||| |  | ||| |||||
Db     115 TCCATGGACCAAGGGAAAGTCTCCTCTGTATTTTATACCATTGTTGTGCCCATGCTA 59
```

```
RESULT  10
BH342817/c
LOCUS       BH342817                649 bp    DNA     linear   GSS 03-DEC-2001
DEFINITION  CH230-64F1.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-64F1, DNA sequence.
ACCESSION   BH342817
VERSION     BH342817.1  GI:17273551
KEYWORDS    GSS.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-64F1.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 64  row: F  column: 1
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Rattus norvegicus"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-64F1"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      219 a    124 c    150 g    156 t
ORIGIN
```

```
  Query Match             26.5%;  Score 129.2;  DB 12;  Length 649;
  Best Local Similarity   55.6%;  Pred. No. 2.4e-20;
  Matches  270;  Conservative    0;  Mismatches  213;  Indels    3;  Gaps    1;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||  ||||| ||||| ||| |||||||  |  || |  ||||||  |  |  || ||
Db     550 TGCTGCCATCTGTAAGCCCCTGCTTTACCAGGTCACCATGTCCCATCAGGTCTGCCACTT 491

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           | || |          |||||||| || ||  |   | |  |||  |||    ||    |
Db     490 GATGATTGTGGCGGTGTATGTAATGGGGTTTGTTGGTGCCATGGCCCATACAGGTAGTGT 431

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||     | || |||||       ||||| ||  || | |     |||||  |   ||
Db     430 TTTAAACATGACCTTCTGTGATGGCAACATCATCAATCACTATATGTGTGATGTACCACC 371

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||     ||| | || ||||   |  | || |  || |  ||  |  | || |||||||
Db     370 TCTCAGGAAACTCTCCTGCACAAGCACCTCCAGCAATGAGCTGGTGGTTTTCATTTTTGT 311

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           | | | |    ||||    ||  | ||||| || |  | ||||| ||     | || |
Db     310 GGGTGTCAATGTAATAATACCTACATTGACTCTCTTTGTTTCTTACACCTTGATCCTTTC 251

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           | || || |  |   | |||| | | ||  |    |||||||||   || || ||| |
Db     250 CAACATCCTCAGCATCCATTCTGCAGAAGGTCGCTCAAAAGCCTTCAGAACCTGTGGCTC 191

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           |||| |     |||| ||  | | || |||| || |   ||||||||||  |    |||
Db     190 CCATGTAGCAGCTGTTTCTCTTTTCTTCGGAGCTGTAGCCTTCATGTATTTTAAGCCT-- 131

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           ||| |   | ||| |||  |||| ||| | | | |  | ||||| || ||| |   |||
Db     132 -TCTAGTGCACCTGTAGATGAAGAGAAAATCTCTACCATTTTTTATACCATTGTGGGTCC 74

Qy     481 cctgct 486
           ||||
Db      73 AATGCT 68
```

```
RESULT  11
BH109232/c
LOCUS       BH109232                763 bp    DNA     linear   GSS 19-JUL-2001
DEFINITION  RPCI-24-352I15.TVB RPCI-24 Mus musculus genomic clone
            RPCI-24-352I15, DNA sequence.
ACCESSION   BH109232
VERSION     BH109232.1  GI:14942107
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-352I15.TJB
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 352  row: I  column: 15
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .763
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-352I15"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
BASE COUNT      265 a    108 c    133 g    257 t
ORIGIN
```

```
  Query Match             26.0%;  Score 126.8;  DB 12;  Length 763;
  Best Local Similarity   58.2%;  Pred. No. 8.8e-20;
  Matches  241;  Conservative    0;  Mismatches  172;  Indels    1;  Gaps    1;

Qy      75 catatgtaattggtttcctgcatcctctggttcatgtgagtttactattgcgactaacttt 134
           | ||  ||   ||   ||| ||| |  ||  |      ||  | ||  | |   | ||||
Db     758 CCTACATAGCAGGGAACCTACATTCCATGAATTCACATAGGGTTCTTGTTCAGGTTACTT 699

Qy     135 tctgcaggtttaacataatacattatttctactgtgaaattttacaactgttcaaaatttt 194
           |||||||||  | || ||||  |  | ||||  |||||  |  | |  ||     || | |
Db     698 TCTGCAGGTCTCACGTAATCAAGCACTTCTTTTGTGATGTCCTTCCCCTAACAGAACTCT 639

Qy     195 catgcaatggtccatc-tattaacgcactaataatatttatttttggtgcttttatacaa 253
           ||||   | | || |  ||| || | ||| || ||| | ||||||  || ||   | |||
Db     638 CATGTGGTTGACCCTTATATCAATGAACTGATGATACTCATTTTTTCTGGTTCACTTCAA 579

Qy     254 atacccacttttaatgactatcataatctcttatactcgtgtgctctttgatattctgaaa 313
           |    ||  | | |  || | || |||||||   |  | || |||  |||  | | |
Db     578 ACTTTTACCATTACTATAGTCCTCATTTCTTATATCTGCATCCTTTTTACTATATTCACA 519

Qy     314 aaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcccatctgctttct 373
           |  |||||   | |||| |||||||||||||| || || || |   ||||  | || |||
Db     518 ATGAAGTCCAGAGAGGGAAGAAGCAAAGCCTTATCTACTTGTGCATCCCACTTTCTGTCT 459

Qy     374 gtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcatctggcttagct 433
           || ||| | | ||| ||  ||||  ||| |||||||  | || ||    ||       |
Db     458 GTGTCAATATTCTATGGGTCTCTCCTCTACATGTATATTCGACCAAGTTCACTTAATGAA 399

Qy     434 gaagaccaagacaaagtgtattctctgttttacacgattataattccccctgcta 487
           |   |  |||||| |     | || | ||||| ||  |  |||||||  |  ||
Db     398 GGGTATAAAGACATACCTGTTGCTATATTTTATACTCTAGTAATTCCTTTATTA 345
```

```
RESULT  12
AZ103967/c
LOCUS       AZ103967                611 bp    DNA     linear   GSS 09-MAY-2000
DEFINITION  RPCI-23-33G18.TV RPCI-23 Mus musculus genomic clone RPCI-23-33G18,
            DNA sequence.
ACCESSION   AZ103967
VERSION     AZ103967.1  GI:7757025
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 611)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-33G18.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 33  row: G  column: 18
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .611
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-33G18"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      206 a    113 c    142 g    150 t
ORIGIN
```

```
  Query Match             25.9%;  Score 126;  DB 12;  Length 611;
  Best Local Similarity   53.7%;  Pred. No. 1.4e-19;
  Matches  261;  Conservative    0;  Mismatches  225;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           || |||||||||||| |||||||| ||  || | || |||||     | |||   |
Db     571 TGCAGCCATATGTAAGCCCTTGCTGTACACACTTATTATGTCTCCACATCTCTCTGTGTT 512

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           |||  | | |    | ||||  |||| | |  |   | |       ||    | |||
Db     511 ATTGGTTATTGGGCCTTATGCCATTGCTCTTATAAGTACAACAACACACACAACTTTGAC 452

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||  |  || | |||||  |     | || ||  |  | || | ||||||  ||||  |
Db     451 CTTTTGCTTACCATTCTGTGGTCCATATATTATCAACCACTTTTTCTGTGACATTTCCCC 392

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
```

```
Db    391 ATTGCTGTCTCTAGCATGTTCTGACACCCACATAAATAAGTTGGTGCTTTTTGTCTTGGC 332

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db    331 TGGCACAGTAGGTGTGCTCAGTGGACTGATAATCCTGGTGTCCTATGTTTGCATCTTGAA 272

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db    271 GGCAATTTTGAAGATTCAGACAGCCAATGGGAGACGAAAAGCCTTCTCAACTTGTTCCTC 212

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db    211 TCACTTGGCAACTGTCTCTATCCTGTATGGGACTCTTTTCTTCATCTATGTTAGACCCAA 152

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db    151 TGTCAGTTCCTCCTTGAATATTAATAAAGTGATCTCTTTATTTTACACGATGGTGATTCC 92

Qy    481 cctgct 486
Db     91 CATGTT 86


RESULT 13
AZ560017/c
LOCUS       AZ560017               507 bp    DNA     linear   GSS 20-NOV-2000
DEFINITION  RPCI-23-216D21.TV RPCI-23 Mus musculus genomic clone RPCI-23-216D21
            , DNA sequence.
ACCESSION   AZ560017
VERSION     AZ560017.1  GI:11239837
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-216D21.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 216  row: D  column: 21
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-216D21"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT      175 a     93 c    134 g    105 t
ORIGIN


  Query Match           25.5%;  Score 124.4;  DB 12;  Length 507;
  Best Local Similarity   53.5%;  Pred. No. 3.2e-19;
  Matches  260;  Conservative    0;  Mismatches  226;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
Db    488 TGTAGCAATCTGTTCACCTCTCCTCTATTCAACACTCATGTCCTCTAGAGTCTGTTTTCT 429

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
Db    428 ATTGTTGGGAATATCCTATGTTGGTGGCTTTGTGAATGGTTGGACATTTACTGGTTGTGT 369

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
Db    368 GTTAAGTCTGTCCTTCTGTGGACCAACTCAGATAAATCACTTTTTCTGTGACTTCTCCCC 309

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
Db    308 TTTGCTGAAAGTGTCCTGCTCAGATGTCTCCATTATTGGAATCATCCCGTCTATCTCTTC 249

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
Db    248 TGGCTCCATTATTGTGGTGACAGTTTTTGTCATAGCTGTCTCCTACATCTACATACTTAT 189

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
Db    188 CACCATCCTGAAGATGCGCTCCACTGAGGGCCGACACAAGGCCTTCTCCACCTGCACCTC 129

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
Db    128 CCACCTCACTGCAGTCACTCTCTTCTATGGGACCATTACCGTCATTTACGTGATGCCCAA 69

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
Db     68 ATCGAGCTACTCTACTGAACAGAACAAGGTGATCTCTTTATTCTACACAGTGGTGATTCC 9

Qy    481 cctgct 486
Db      8 TATGTT 3


RESULT 14
AZ086388/c
LOCUS       AZ086388               628 bp    DNA     linear   GSS 08-MAY-2000
DEFINITION  RPCI-23-26C8.TV RPCI-23 Mus musculus genomic clone RPCI-23-26C8,
            DNA sequence.
ACCESSION   AZ086388
VERSION     AZ086388.1  GI:7728122
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 628)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-26C8.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
```

```
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 26  row: C  column: 8
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .628
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-26C8"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      211 a    128 c    164 g    125 t
ORIGIN


  Query Match           25.5%;  Score 124.4;  DB 12;  Length 628;
  Best Local Similarity   53.5%;  Pred. No. 3.2e-19;
  Matches  260;  Conservative    0;  Mismatches  226;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          |||||| || |||   ||  | || ||| ||    | ||||||   | | || |   ||
Db    488 TGTAGCAATCTGTTCACCTCTCCTCTATTCAACACTCATGTCCTCTAGAGTCTGTTTTCT 429

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           ||| |  | || || |||||   ||| ||  || ||  |  |     |    |||   |
Db    428 ATTGTTGGGAATATCCTATGTTGGTGGCTTTGTGAATGGTTGGACATTTACTGGTTGTGT 369

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||  |  ||  | |||||  |     |     ||| || | || | ||||||  | |  |
Db    368 GTTAAGTCTGTCCTTCTGTGGACCAACTCAGATAAATCACTTTTTCTGTGACTTCTCCCC 309

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           || | ||| | || |||   | |   || ||||  | | | ||    | ||| | |
Db    308 TTTGCTGAAAGTGTCCTGCTCAGATGTCTCCATTATTGGAATCATCCCGTCTATCTCTTC 249

Qy    241 tgcttttatacaaatacccacttaatgactatcataatctcttatactcgtgtgctctt 300
          ||  |  ||     |    ||  |  |    ||     |||| || |      | ||  |
Db    248 TGGCTCCATTATTGTGGTGACAGTTTTTGTCATAGCTGTCTCCTACATCTACATACTTAT 189

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
            || ||||| |     ||     ||||| ||  ||| ||||||||||| |||  | |
Db    188 CACCATCCTGAAGATGCGCTCCACTGAGGGCCGACACAAGGCCTTCTCCACCTGCACCTC 129

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          |||  ||   | | ||| |  | | ||| || ||  | | | |||| || |||   ||
Db    128 CCACCTCACTGCAGTCACTCTCTTCTATGGGACCATTACCGTCATTTACGTGATGCCCAA 69

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          |||  |||   ||   || ||  |||| |||   ||| | || |||||  |  | |||||
Db     68 ATCGAGCTACTCTACTGAACAGAACAAGGTGATCTCTTTATTCTACACAGTGGTGATTCC 9

Qy    481 cctgct 486
            || |
Db      8 TATGTT 3


RESULT  15
AZ607393/c
LOCUS       AZ607393                740 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0429N15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0429N15 R, DNA sequence.
ACCESSION   AZ607393
VERSION     AZ607393.1  GI:11729583
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0429  row: N  column: 15
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1. .740
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0429N15"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      276 a    134 c    159 g    171 t
ORIGIN


  Query Match           24.3%;  Score 118.2;  DB 12;  Length 740;
  Best Local Similarity   55.5%;  Pred. No. 9.2e-18;
  Matches  248;  Conservative    0;  Mismatches  198;  Indels    1;  Gaps    1;

Qy      5 gccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcagttg 64
          |||||  |||| ||| |  | ||| ||||    |||||||  ||| |  ||   || ||
Db    448 GCCATTAGTAACCCCCTTTTGTATGCAGTAGACATGTCCAGGAAAGTGTGCTTCCAATTA 389

Qy     65 ctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttactattg 124
            | | |  |   |||    | |  ||      || || || | |||   |  ||   |||
Db    388 TTGACTGGTGTTTATTCTGTGGCATTAGCAGATGCTTTGATACATACAACATTGACATTC 329
```

```
Qy    125 cgactaactttctgcaggtttaacataatacattatttctactgtgaaattttacaactg 184
          |   ||  ||||||  |||  ||    ||  || ||||||  ||||||  |||   | | ||
Db    328 CACTTATGTTTCTGTGGGTCGAATGAGATTAATCATTTCTTCTGTGATATTCCTCCAGTG 269

Qy    185 ttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggtgct 244
          ||   | | || ||    |   ||    | || |   || | ||||| | | ||  || |
Db    268 TTAGTACTGTCTTGTTCAGACACACAGGTCAATGTTTTAGTCATATTCACTGTTTTTGGT 209

Qy    245 tttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctttgat 304
          || ||  || |   |||  |        |  | ||||| ||   |    |  |||
Db    208 TTCATTGAACTGAGCACCATCTCAGGAGTTCTTATCTCCTACTGTTACATCATCTCATCA 149

Qy    305 attctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcccat 364
           |||||| |  |  ||||    || ||    |||||  ||||| || ||     | ||.
Db    148 GTTCTGAAGATCAGTTCTGCTGCTGGGAGGTTAAAAGCTTTCTCTACCTGTACATCACAC 89

Qy    365 ctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgcatct 424
          |||  | ||||  || | |   | ||||||  || ||||||||||| | || ||    |||
Db     88 CTGACTGCTGTTGCAATTTTTCAAGGAACTATGCTCTTCATGTAT-TCCGGCCAAGTTCT 30

Qy    425 ggcttagctgaagaccaagacaaagtg 451
            ||   |   ||| ||||| ||||||
Db     29 TCCTACTCCCTAGATCAAGATAAAGTG 3
```

Search completed: June 27, 2002, 10:45:49
Job time: 1642 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 11:17:12 ; Search time 1801.95 Seconds
(without alignments)
5655.664 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	487	100.0	487	6	AX181436	AX181436 Sequence
2	487	100.0	487	9	AF179759	AF179759 Homo sapi
3	486	99.8	486	6	AX242110	AX242110 Sequence
4	236	48.5	485	6	AX181383	AX181383 Sequence
5	236	48.5	485	9	AF179729	AF179729 Pan trogl
6	185	38.0	487	6	AX181394	AX181394 Sequence
7	185	38.0	487	9	AF179735	AF179735 Pan trogl
8	30	6.2	927	10	AY073094	AY073094 Mus muscu
9	30	6.2	138556	2	AC105564	AC105564 Rattus no
10	29	6.0	988	6	AX241993	AX241993 Sequence
11	29	6.0	150083	2	AC024892	AC024892 Homo sapi
12	26	5.3	919	10	AY074074	AY074074 Mus muscu
13	26	5.3	930	10	AY073315	AY073315 Mus muscu
14	23	4.7	71291	9	AC090115	AC090115 Homo sapi
c 15	23	4.7	138556	2	AC105564	AC105564 Rattus no
16	22	4.5	156957	9	AP001977	AP001977 Homo sapi
c 17	22	4.5	160800	2	AP002011	AP002011 Homo sapi
18	22	4.5	183674	2	AC016996	AC016996 Homo sapi
19	21	4.3	2014	5	AF083078	AF083078 Carassius
20	21	4.3	155183	2	AC032016	AC032016 Homo sapi
21	21	4.3	161699	9	AL445217	AL445217 Human DNA
c 22	21	4.3	164173	2	AC011848	AC011848 Homo sapi
23	21	4.3	164679	8	AC078944	AC078944 Genomic S
24	20	4.1	4381	2	AC012953	AC012953 Drosophil
c 25	20	4.1	29963	9	AL356859	AL356859 Human DNA
26	20	4.1	40639	9	AL359535	AL359535 Human DNA
c 27	20	4.1	40837	3	AF067947	AF067947 Caenorhab
28	20	4.1	71704	2	AC068468	AC068468 Homo sapi
c 29	20	4.1	71704	2	AC068468	AC068468 Homo sapi
c 30	20	4.1	84591	9	AC096655	AC096655 Homo sapi
31	20	4.1	95794	2	AC108568	AC108568 Rattus no
32	20	4.1	110000	2	AC013623_2	Continuation (3 of
c 33	20	4.1	111561	2	AC098034	AC098034 Rattus no
34	20	4.1	128485	2	AL358814	AL358814 Homo sapi
c 35	20	4.1	134982	8	AP003279	AP003279 Oryza sat
c 36	20	4.1	150167	9	AC034150	AC034150 Homo sapi
c 37	20	4.1	150442	2	AC022782	AC022782 Mus muscu
38	20	4.1	159453	2	AC006732	AC006732 Caenorhab
c 39	20	4.1	167596	30	AC016137	Ac016137 Homo sapi
40	20	4.1	168965	2	AC105751	AC105751 Homo sapi
c 41	20	4.1	169112	3	AC023716	AC023716 Drosophil
42	20	4.1	169234	9	CNS06C7Q	AL390799 Human chr
c 43	20	4.1	170288	3	AC023694	AC023694 Drosophil
c 44	20	4.1	192178	2	AC009443	AC009443 Homo sapi
45	20	4.1	193534	2	AC013464	AC013464 Homo sapi

ALIGNMENTS

RESULT 1
AX181436
LOCUS AX181436 487 bp DNA linear PAT 07-AUG-2001
DEFINITION Sequence 224 from Patent WO0146262.
ACCESSION AX181436
VERSION AX181436.1 GI:15141548
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 487)
AUTHORS Rouquier,S. and Giorgi,D.
TITLE Olfactory receptor genes and pseudogenes in primates and mouse
JOURNAL Patent: WO 0146262-A 224 28-JUN-2001;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES Location/Qualifiers
source 1. .487
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Taxon = 9606; gene = HSA1; Accession
DDBJ/EMBL/GenBank = AF179759"
CDS <2. .>487

```
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="CAC50218.1"
                     /db_xref="GI:15141549"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDI
                     LKKKSEKGRSKAFSTCGAHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIP
                     LL"
BASE COUNT      131 a    105 c     77 g    174 t
ORIGIN


  Query Match             100.0%;  Score 487;  DB 6;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-251;
  Matches  487;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGCTAAGTATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATAATATTTATTTTTGG 240

Qy     241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGCTTTTATACAAATACCCACTTTAATGACTATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCGGCGC 360

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATCTGGCTTAGCTGAAGACCAAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCC 480

Qy     481 cctgcta 487
           |||||||
Db     481 CCTGCTA 487


RESULT   2
AF179759
LOCUS       AF179759                 487 bp    DNA     linear   PRI 09-MAR-2000
DEFINITION  Homo sapiens olfactory receptor (HSA1) gene, partial cds.
ACCESSION   AF179759
VERSION     AF179759.1  GI:7211526
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Rouquier,S., Blancher,A. and Giorgi,D.
  TITLE     The olfactory receptor gene repertoire in primates and mouse:
            evidence for reduction of the functional fraction in primates
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
  MEDLINE   20183981
REFERENCE   2  (bases 1 to 487)
  AUTHORS   Giorgi,D.G. and Rouquier,S.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
            1142, rue de la Cardonille, Montpellier Cedex 5 34396, France
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            <1. .>487
                     /gene="HSA1"
     CDS             <1. .>487
                     /gene="HSA1"
                     /codon_start=2
                     /product="olfactory receptor"
                     /protein_id="AAF40348.1"
                     /db_xref="GI:7211527"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALIIFIFGAFIQIPTLMTIIISYTRVLFDI
                     LKKKSEKGRSKAFSTCGAHLLSVSLYYGTLIFMYVRPASGLAEDQDKVYSLFYTIIIP
                     LL"
BASE COUNT      131 a    105 c     77 g    174 t
ORIGIN


  Query Match             100.0%;  Score 487;  DB 9;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-251;
  Matches  487;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGCTAAGTATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATAATATTTATTTTTGG 240

Qy     241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGCTTTTATACAAATACCCACTTTAATGACTATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCGGCGC 360

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy     421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATCTGGCTTAGCTGAAGACCAAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCC 480

Qy     481 cctgcta 487
           |||||||
Db     481 CCTGCTA 487


RESULT   3
AX242110
LOCUS       AX242110                 486 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 858 from Patent WO0127158.
ACCESSION   AX242110
VERSION     AX242110.1  GI:15798985
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
```

```
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 858 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g708 nucleotide)"
BASE COUNT      131 a    105 c     77 g    173 t
ORIGIN


  Query Match             99.8%;  Score 486;  DB 6;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-250;
  Matches  486;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCAG 60

Qy     62 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTGCTAAGTATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACTA 120

Qy    122 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttacaa 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACAA 180

Qy    182 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATAATATTTATTTTTGGT 240

Qy    242 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCTTTTATACAAATACCCACTTTAATGACTATCATAATCTCTTATACTCGTGTGCTCTTT 300

Qy    302 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCGGCGCC 360

Qy    362 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGCA 420

Qy    422 tctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattccc 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCTGGCTTAGCTGAAGACCAAGACAAAGTGTATTCTCTGTTTTACACGATTATAATTCCC 480

Qy    482 ctgcta 487
          ||||||
Db    481 CTGCTA 486


RESULT   4
AX181383
LOCUS       AX181383                485 bp    DNA     linear   PAT 07-AUG-2001
DEFINITION  Sequence 171 from Patent WO0146262.
ACCESSION   AX181383
VERSION     AX181383.1  GI:15132985
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Rouquier,S. and Giorgi,D.
  TITLE     Olfactory receptor genes and pseudogenes in primates and mouse
  JOURNAL   Patent: WO 0146262-A 171 28-JUN-2001;
            CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     misc_feature    1. .485
                     /note="Taxon = 9598; gene = PTR204; pseudogene; Accession
                     DDBJ/EMBL/GenB a"
BASE COUNT      130 a    107 c     77 g    171 t
ORIGIN


  Query Match             48.5%;  Score 236;  DB 6;  Length 485;
  Best Local Similarity   99.1%;  Pred. No. 7.3e-116;
  Matches  436;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy    241 tgctttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagacc 440
          ||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACC 440


RESULT   5
AF179729
LOCUS       AF179729                485 bp    DNA     linear   PRI 09-MAR-2000
DEFINITION  Pan troglodytes PTR204 pseudogene, partial sequence.
ACCESSION   AF179729
VERSION     AF179729.1  GI:7211473
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Rouquier,S., Blancher,A. and Giorgi,D.
  TITLE     The olfactory receptor gene repertoire in primates and mouse:
            evidence for reduction of the functional fraction in primates
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
  MEDLINE   20183981
REFERENCE   2  (bases 1 to 485)
  AUTHORS   Giorgi,D.G. and Rouquier,S.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
            1142, rue de la Cardonille, Montpellier Cedex 5 34396, France
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
```

```
     gene            <1. .>485
                     /gene="PTR204"
                     /pseudo
BASE COUNT      130 a    107 c     77 g    171 t
ORIGIN
```

```
  Query Match             48.5%;  Score 236;  DB 9;  Length 485;
  Best Local Similarity   99.1%;  Pred. No. 7.3e-116;
  Matches  436;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATACTCGTGTGCTCTT 300

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy     421 atctggcttagctgaagacc 440
           ||||||||||||||||||||
Db     421 ATCTGGCTTAGCTGAAGACC 440
```

```
RESULT   6
AX181394
LOCUS       AX181394                 487 bp    DNA     linear   PAT 07-AUG-2001
DEFINITION  Sequence 182 from Patent WO0146262.
ACCESSION   AX181394
VERSION     AX181394.1  GI:15141512
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Rouquier,S. and Giorgi,D.
  TITLE     Olfactory receptor genes and pseudogenes in primates and mouse
  JOURNAL   Patent: WO 0146262-A 182 28-JUN-2001;
            CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /note="Taxon = 9598; gene = PTR210; Accession
                     DDBJ/EMBL/GenBank = AF1797"
     CDS             <2. .>487
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="CAC50200.1"
                     /db_xref="GI:15141513"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALMIFIFGAFIQIPTLMTIIISYSRVLFDI
```

```
                     LKKKSEKGRSKAFSTCSAHLLSVSLYYGTLIFMYVRPASGLAEDPDKVYSLFYTIIIP
                     LL"
BASE COUNT      129 a    107 c     78 g    173 t
ORIGIN
```

```
  Query Match             38.0%;  Score 185;  DB 6;  Length 487;
  Best Local Similarity   98.9%;  Pred. No. 2.1e-88;
  Matches  435;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy      61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy     121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy     181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy     241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATTCTCGTGTGCTCTT 300

Qy     301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy     361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy     421 atctggcttagctgaagacc 440
           ||||||||||||||||||||
Db     421 ATCTGGCTTAGCTGAAGACC 440
```

```
RESULT   7
AF179735
LOCUS       AF179735                 487 bp    DNA     linear   PRI 09-MAR-2000
DEFINITION  Pan troglodytes olfactory receptor (PTR210) gene, partial cds.
ACCESSION   AF179735
VERSION     AF179735.1  GI:7211484
KEYWORDS    .
SOURCE      chimpanzee.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Rouquier,S., Blancher,A. and Giorgi,D.
  TITLE     The olfactory receptor gene repertoire in primates and mouse:
            evidence for reduction of the functional fraction in primates
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
  MEDLINE   20183981
REFERENCE   2  (bases 1 to 487)
  AUTHORS   Giorgi,D.G. and Rouquier,S.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
            1142, rue de la Cardonille, Montpellier Cedex 5 34396, France
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     gene            <1. .>487
                     /gene="PTR210"
     CDS             <1. .>487
                     /gene="PTR210"
```

```
                     /codon_start=2
                     /product="olfactory receptor"
                     /protein_id="AAF40330.1"
                     /db_xref="GI:7211485"
                     /translation="VAICNPLLYPVMMSNKLSAQLLSISYVIGFLHPLVHVSLLLRLT
                     FCRFNIIHYFYCEILQLFKISCNGPSINALMIFIFGAFIQIPTLMTIIISYSRVLFDI
                     LKKKSEKGRSKAFSTCSAHLLSVSLYYGTLIFMYVRPASGLAEDPDKVYSLFYTIIIP
                     LL"
BASE COUNT      129 a    107 c     78 g    173 t
ORIGIN


  Query Match           38.0%;  Score 185;  DB 9;  Length 487;
  Best Local Similarity   98.9%;  Pred. No. 2.1e-88;
  Matches  435;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTAGCCATATGTAATCCCTTGCTTTATCCAGTGATGATGTCCAACAAACTCAGCGCTCA 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCTAAGCATTTCATATGTAATTGGTTTCCTGCATCCTCTGGTTCATGTGAGTTTACT 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGCGACTAACTTTCTGCAGGTTTAACATAATACATTATTTCTACTGTGAAATTTTACA 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    181 ACTGTTCAAAATTTCATGCAATGGTCCATCTATTAACGCACTAATGATATTTATTTTTGG 240

Qy    241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db    241 TGCTTTTATACAAATACCCACTTTAATGACGATCATAATCTCTTATTCTCGTGTGCTCTT 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    301 TGATATTCTGAAAAAAAAGTCTGAAAAGGGCAGAAGCAAAGCCTTCTCCACATGCAGCGC 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCATCTGCTTTCTGTCTCATTGTACTACGGAACTCTGATCTTCATGTATGTGCGTCCTGC 420

Qy    421 atctggcttagctgaagacc 440
          ||||||||||||||||||||
Db    421 ATCTGGCTTAGCTGAAGACC 440


RESULT   8
AY073094
LOCUS       AY073094                 927 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR182-2 gene, complete cds.
ACCESSION   AY073094
VERSION     AY073094.1  GI:18479485
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 927)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
   PUBMED   11802173
REFERENCE   2  (bases 1 to 927)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .927
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     mRNA            <1. .>927
                     /product="olfactory receptor MOR182-2"
     CDS             1. .927
                     /codon_start=1
                     /evidence=not_experimental
                     /product="olfactory receptor MOR182-2"
                     /protein_id="AAL60757.1"
                     /db_xref="GI:18479486"
                     /translation="MAEGNRTLVTEFILMGLTDHPTLKAALFPLFLVIYLITMVGNLG
                     LIALIWKDSHLHTPMYLFLSSLAFADSCTTSSVTPRMILNFLSTNHEITLVECFAQFY
                     FMGSSATTECFLLSVMAYDRYLAICNPLLYPVLMSNRLCTQFIAVAYLLGALHLAIHV
                     GLLLRLTFCRSNIIQYYYCEILQLFNISCIDPTINVFVLLILATSIQVFTFVTVLVSY
                     IRVLFAILRKKSDKGRSKAFSTCSAHLLSVSLFYGTLFLIYVCPGSGPVGDKEKMLSL
                     FYTVIIPLLNPFVYSLRNKEVISAFRRVMKNT"
BASE COUNT      221 a    229 c    166 g    311 t
ORIGIN


  Query Match            6.2%;  Score 30;  DB 10;  Length 927;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    326 aagggcagaagcaaagccttctccacatgc 355
          ||||||||||||||||||||||||||||||
Db    694 AAGGGCAGAAGCAAAGCCTTCTCCACATGC 723


RESULT   9
AC105564
LOCUS       AC105564               138556 bp    DNA     linear   HTG 09-JAN-2002
DEFINITION  Rattus norvegicus clone CH230-250G11, *** SEQUENCING IN PROGRESS
            ***, 60 unordered pieces.
ACCESSION   AC105564
VERSION     AC105564.1  GI:18092786
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 138556)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
```

Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 138556)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT ---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GNEZ
Center clone name: CH230-250G11
---------- Summary Statistics
Assembly program: Phrap; version 0.990329First call to findPhrapList
Consensus quality: 104982 bases at least Q40
Consensus quality: 113437 bases at least Q30
Consensus quality: 120052 bases at least Q20
Estimated insert size: 98776; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.3x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	12026	contig of 12026 bp in length
*	12027	12126	gap of unknown length
*	12127	17583	contig of 5457 bp in length
*	17584	17683	gap of unknown length
*	17684	22927	contig of 5244 bp in length
*	22928	23027	gap of unknown length
*	23028	27987	contig of 4960 bp in length
*	27988	28087	gap of unknown length
*	28088	32504	contig of 4417 bp in length
*	32505	32604	gap of unknown length
*	32605	35333	contig of 2729 bp in length
*	35334	35433	gap of unknown length
*	35434	37327	contig of 1894 bp in length
*	37328	37427	gap of unknown length
*	37428	40912	contig of 3485 bp in length
*	40913	41012	gap of unknown length
*	41013	44129	contig of 3117 bp in length
*	44130	44229	gap of unknown length
*	44230	46597	contig of 2368 bp in length
*	46598	46697	gap of unknown length
*	46698	50418	contig of 3721 bp in length
*	50419	50518	gap of unknown length
*	50519	52990	contig of 2472 bp in length
*	52991	53090	gap of unknown length
*	53091	56522	contig of 3432 bp in length
*	56523	56622	gap of unknown length
*	56623	59005	contig of 2383 bp in length
*	59006	59105	gap of unknown length
*	59106	61667	contig of 2562 bp in length
*	61668	61767	gap of unknown length
*	61768	63899	contig of 2132 bp in length
*	63900	63999	gap of unknown length
*	64000	68087	contig of 4088 bp in length
*	68088	68187	gap of unknown length
*	68188	70694	contig of 2507 bp in length
*	70695	70794	gap of unknown length
*	70795	73386	contig of 2592 bp in length
*	73387	73486	gap of unknown length
*	73487	75692	contig of 2206 bp in length
*	75693	75792	gap of unknown length
*	75793	77929	contig of 2137 bp in length
*	77930	78029	gap of unknown length
*	78030	79296	contig of 1267 bp in length
*	79297	79396	gap of unknown length
*	79397	82238	contig of 2842 bp in length
*	82239	82338	gap of unknown length
*	82339	84658	contig of 2320 bp in length
*	84659	84758	gap of unknown length
*	84759	87336	contig of 2578 bp in length
*	87337	87436	gap of unknown length
*	87437	89496	contig of 2060 bp in length
*	89497	89596	gap of unknown length
*	89597	90917	contig of 1321 bp in length
*	90918	91017	gap of unknown length
*	91018	92625	contig of 1608 bp in length
*	92626	92725	gap of unknown length
*	92726	93955	contig of 1230 bp in length
*	93956	94055	gap of unknown length
*	94056	95238	contig of 1183 bp in length
*	95239	95338	gap of unknown length
*	95339	96506	contig of 1168 bp in length
*	96507	96606	gap of unknown length
*	96607	98414	contig of 1808 bp in length
*	98415	98514	gap of unknown length
*	98515	100270	contig of 1756 bp in length
*	100271	100370	gap of unknown length
*	100371	101914	contig of 1544 bp in length
*	101915	102014	gap of unknown length
*	102015	103378	contig of 1364 bp in length
*	103379	103478	gap of unknown length
*	103479	104660	contig of 1182 bp in length
*	104661	104760	gap of unknown length
*	104761	106384	contig of 1624 bp in length
*	106385	106484	gap of unknown length
*	106485	107638	contig of 1154 bp in length
*	107639	107738	gap of unknown length
*	107739	109065	contig of 1327 bp in length
*	109066	109165	gap of unknown length
*	109166	110732	contig of 1567 bp in length
*	110733	110832	gap of unknown length
*	110833	112101	contig of 1269 bp in length
*	112102	112201	gap of unknown length
*	112202	113509	contig of 1308 bp in length
*	113510	113609	gap of unknown length
*	113610	115100	contig of 1491 bp in length
*	115101	115200	gap of unknown length
*	115201	116535	contig of 1335 bp in length
*	116536	116635	gap of unknown length
*	116636	118099	contig of 1464 bp in length
*	118100	118199	gap of unknown length
*	118200	119225	contig of 1026 bp in length
*	119226	119325	gap of unknown length
*	119326	121099	contig of 1774 bp in length
*	121100	121199	gap of unknown length
*	121200	122556	contig of 1357 bp in length
*	122557	122656	gap of unknown length
*	122657	124309	contig of 1653 bp in length
*	124310	124409	gap of unknown length
*	124410	125647	contig of 1238 bp in length
*	125648	125747	gap of unknown length
*	125748	126786	contig of 1039 bp in length

```
            *   126787   126886: gap of unknown length
            *   126887   128160: contig of 1274 bp in length
            *   128161   128260: gap of unknown length
            *   128261   129777: contig of 1517 bp in length
            *   129778   129877: gap of unknown length
            *   129878   130976: contig of 1099 bp in length
            *   130977   131076: gap of unknown length
            *   131077   132445: contig of 1369 bp in length
            *   132446   132545: gap of unknown length
            *   132546   133844: contig of 1299 bp in length
            *   133845   133944: gap of unknown length
            *   133945   135012: contig of 1068 bp in length

  Query Match             6.2%;  Score 30;  DB 2;  Length 138556;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-05;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    326 aagggcagaagcaaagccttctccacatgc 355
          ||||||||||||||||||||||||||||||
Db 136474 AAGGGCAGAAGCAAAGCCTTCTCCACATGC 136503


RESULT  10
AX241993
LOCUS       AX241993                 988 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 741 from Patent WO0127158.
ACCESSION   AX241993
VERSION     AX241993.1  GI:15798868
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 988)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 741 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .988
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g590 nucleotide)"
BASE COUNT      250 a    231 c    165 g    342 t
ORIGIN


  Query Match             6.0%;  Score 29;  DB 6;  Length 988;
  Best Local Similarity   100.0%;  Pred. No. 0.00021;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    152 atacattatttctactgtgaaatttaca 180
          |||||||||||||||||||||||||||||
Db    521 ATACATTATTTCTACTGTGAAATTTTACA 549


RESULT  11
AC024892
LOCUS       AC024892               150083 bp    DNA     linear   HTG 15-OCT-2001
DEFINITION  Homo sapiens chromosome 3 clone RP11-214N20, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC024892
VERSION     AC024892.11  GI:16117893
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 150083)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150083)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Oct 14, 2001 this sequence version replaced gi:9438337.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HAIU
                Center clone name: RP11-214N20
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-primer Bodipy: 2% of reads
                Chemistry: Dye-terminator Big Dye: 98% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 137907 bases at least Q40
                Consensus quality: 146103 bases at least Q30
                Consensus quality: 149133 bases at least Q20
                Estimated insert size: 151155; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 4.1x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
```

```
            * be preserved.
            *       1    34152: contig of 34152 bp in length
            *   34153    34252: gap of unknown length
            *   34253    55507: contig of 21255 bp in length
            *   55508    55607: gap of unknown length
            *   55608    79422: contig of 23815 bp in length
            *   79423    79522: gap of unknown length
            *   79523    94794: contig of 15272 bp in length
            *   94795    94894: gap of unknown length
            *   94895   107021: contig of 12127 bp in length
            *  107022   107121: gap of unknown length
            *  107122   116733: contig of 9612 bp in length
            *  116734   116833: gap of unknown length
            *  116834   125530: contig of 8697 bp in length
            *  125531   125630: gap of unknown length
            *  125631   135300: contig of 9670 bp in length
            *  135301   135400: gap of unknown length
            *  135401   140864: contig of 5464 bp in length
            *  140865   140964: gap of unknown length
            *  140965   146546: contig of 5582 bp in length
            *  146547   146646: gap of unknown length
            *  146647   150083: contig of 3437 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150083
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-214N20"
BASE COUNT    45301 a  29114 c  27778 g  46869 t   1021 others
ORIGIN


  Query Match             6.0%;  Score 29;  DB 2;  Length 150083;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     152 atacattatttctactgtgaaatttaca 180
           ||||||||||||||||||||||||||||
Db   57737 ATACATTATTTCTACTGTGAAATTTACA 57765


RESULT  12
AY074074
LOCUS       AY074074                 919 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR182-8 pseudogene, partial
            sequence.
ACCESSION   AY074074
VERSION     AY074074.1  GI:18481236
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 919)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
   PUBMED   11802173
REFERENCE   2  (bases 1 to 919)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .919
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     gene            <1. .>919
                     /gene="olfactory receptor MOR182-8"
                     /note="potential coding region disrupted by one or more
                     internal stop codons"
                     /pseudo
                     /evidence=not_experimental
BASE COUNT      223 a    218 c    172 g    306 t
ORIGIN


  Query Match             5.3%;  Score 26;  DB 10;  Length 919;
  Best Local Similarity   100.0%;  Pred. No. 0.0086;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      20 ttgctttatccagtgatgatgtccaa 45
           |||||||||||||||||||||||||||
Db     389 TTGCTTTATCCAGTGATGATGTCCAA 414


RESULT  13
AY073315
LOCUS       AY073315                 930 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR183-1 gene, complete cds.
ACCESSION   AY073315
VERSION     AY073315.1  GI:18479927
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
   PUBMED   11802173
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     mRNA            <1. .>930
                     /product="olfactory receptor MOR183-1"
     CDS             1. .930
                     /codon_start=1
                     /evidence=not_experimental
                     /product="olfactory receptor MOR183-1"
                     /protein_id="AAL60978.1"
                     /db_xref="GI:18479928"
                     /translation="MGIENTTLLTEFVLTGLSHLPQWKIPLFLLFLVIYLITIVGNLG
                     LITLIWNDPHLHIPMYFFLGSLAFVDTWLSSTVTPKMLLDIFSKSKLISFSECMIQFF
                     SFGISATTECFLLASMAYDRYVAICKPLLYPVIMTNRLCVRLLTLSFVGGFIHVLIHE
                     SFLFRLTFCNSNIIHHFYCDVMPLLKISCNDPSLNYLMLFIFSGSIQVFSILTILISY
                     TLVLFSILKQKSLKSIKKAFSTCGAHLLSVSLYYGSLLFMYVRPASPQVDDQDMMDSI
                     FYTVIIPVLNPIIYSLRNKQVKNSLEKFLKRNT"
BASE COUNT      249 a    201 c    143 g    337 t
ORIGIN


  Query Match             5.3%;  Score 26;  DB 10;  Length 930;
  Best Local Similarity   100.0%;  Pred. No. 0.0086;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     399 tcttcatgtatgtgcgtcctgcatct 424
           ||||||||||||||||||||||||||
Db     767 TCTTCATGTATGTGCGTCCTGCATCT 792


RESULT  14
AC090115
LOCUS       AC090115                 71291 bp    DNA     linear   PRI 29-SEP-2001
DEFINITION  Homo sapiens 12q BAC RP11-762H13 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC090115
```

VERSION AC090115.15 GI:15808446
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 71291)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 71291)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 71291)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Sep 29, 2001 this sequence version replaced gi:15559142.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1. .71291
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="12q"
	/clone="RP11-762H13"
misc_feature	1. .2007
	/note="overlaps bases 162939. .164945 of clone AC089987"
	/function="clone overlap"
repeat_region	150. .311
	/rpt_family="L1MD2"
repeat_region	complement(354. .945)
	/rpt_family="MLT1E2"
repeat_region	complement(1230. .1395)
	/rpt_family="MLT1A2"
repeat_region	1396. .1434
	/rpt_family="(CA)n"
repeat_region	complement(1435. .1490)
	/rpt_family="MLT1A2"
repeat_region	1510. .1695
	/rpt_family="L1MEc"
repeat_region	complement(3462. .3756)
	/rpt_family="AluY"
repeat_region	5101. .5407
	/rpt_family="L2"
repeat_region	complement(7173. .7307)
	/rpt_family="AluSg/x"
repeat_region	complement(7492. .7614)
	/rpt_family="MIR"
repeat_region	complement(8407. .8704)
	/rpt_family="AluSx"
repeat_region	8708. .8745
	/rpt_family="AT_rich"
repeat_region	complement(8746. .14867)
	/rpt_family="L1PA4"
STS	10108. .10286
	/standard_name="61535"
STS	10148. .10429
	/standard_name="152451"
repeat_region	14901. .15181
	/rpt_family="AluY"
repeat_region	15347. .15486
	/rpt_family="HERVL"
repeat_region	complement(15499. .15751)
	/rpt_family="L1M4"

```
     repeat_region   complement(15752. .16059)
                     /rpt_family="AluSp"
     repeat_region   complement(16060. .16377)
                     /rpt_family="L1M4"
     repeat_region   complement(16425. .16606)
                     /rpt_family="L1M4"
     repeat_region   complement(16657. .16898)
                     /rpt_family="L1MB8"
     repeat_region   complement(16924. .17196)
                     /rpt_family="L1M4"
     repeat_region   17303. .17330
                     /rpt_family="AT_rich"
     repeat_region   complement(17352. .21082)
                     /rpt_family="L1PB1"
     repeat_region   21083. .23565
                     /rpt_family="L1PA13"
     repeat_region   complement(23594. .23876)
                     /rpt_family="L1ME1"
     repeat_region   complement(23897. .24139)
                     /rpt_family="L1ME1"
     repeat_region   24136. .24981
                     /rpt_family="L1M4"
     repeat_region   26818. .27134
                     /rpt_family="AluSx"
     repeat_region   27148. .27549
                     /rpt_family="MSTA"
     repeat_region   complement(27779. .28324)
                     /rpt_family="MER67D"
     repeat_region   complement(29282. .29364)
                     /rpt_family="MIR"
     repeat_region   29799. .29840
                     /rpt_family="AT_rich"
     repeat_region   29955. .30172
                     /rpt_family="L2"
     repeat_region   30277. .30581
                     /rpt_family="L2"
     repeat_region   30628. .31063
                     /rpt_family="L1PA16"
     repeat_region   complement(31497. .31774)
                     /rpt_family="MSTC"
     repeat_region   complement(31775. .32067)
                     /rpt_family="AluSx"
     repeat_region   complement(32068. .32155)
                     /rpt_family="MSTC"
     repeat_region   complement(32157. .33408)
                     /rpt_family="L2"
     repeat_region   35570. .35875
                     /rpt_family="AluSq"
     repeat_region   36055. .36084
                     /rpt_family="AT_rich"

  Query Match             4.7%;  Score 23;  DB 9;  Length 71291;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     329 ggcagaagcaaagccttctccac 351
           |||||||||||||||||||||||
Db    6390 GGCAGAAGCAAAGCCTTCTCCAC 6412


RESULT  15
AC105564/c
LOCUS       AC105564              138556 bp    DNA     linear   HTG 09-JAN-2002
DEFINITION  Rattus norvegicus clone CH230-250G11, *** SEQUENCING IN PROGRESS
            ***, 60 unordered pieces.
ACCESSION   AC105564
VERSION     AC105564.1  GI:18092786
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE   1  (bases 1 to 138556)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 138556)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GNEZ
                Center clone name: CH230-250G11
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329First call to
            findPhrapList
                Consensus quality: 104982 bases at least Q40
                Consensus quality: 113437 bases at least Q30
                Consensus quality: 120052 bases at least Q20
                Estimated insert size: 98776; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 1.3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 60 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
```

```
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1    12026: contig of 12026 bp in length
*   12027    12126: gap of unknown length
*   12127    17583: contig of 5457 bp in length
*   17584    17683: gap of unknown length
*   17684    22927: contig of 5244 bp in length
*   22928    23027: gap of unknown length
*   23028    27987: contig of 4960 bp in length
*   27988    28087: gap of unknown length
*   28088    32504: contig of 4417 bp in length
*   32505    32604: gap of unknown length
*   32605    35333: contig of 2729 bp in length
*   35334    35433: gap of unknown length
*   35434    37327: contig of 1894 bp in length
*   37328    37427: gap of unknown length
*   37428    40912: contig of 3485 bp in length
*   40913    41012: gap of unknown length
*   41013    44129: contig of 3117 bp in length
*   44130    44229: gap of unknown length
*   44230    46597: contig of 2368 bp in length
*   46598    46697: gap of unknown length
*   46698    50418: contig of 3721 bp in length
*   50419    50518: gap of unknown length
*   50519    52990: contig of 2472 bp in length
*   52991    53090: gap of unknown length
*   53091    56522: contig of 3432 bp in length
*   56523    56622: gap of unknown length
*   56623    59005: contig of 2383 bp in length
*   59006    59105: gap of unknown length
*   59106    61667: contig of 2562 bp in length
*   61668    61767: gap of unknown length
*   61768    63899: contig of 2132 bp in length
*   63900    63999: gap of unknown length
*   64000    68087: contig of 4088 bp in length
*   68088    68187: gap of unknown length
*   68188    70694: contig of 2507 bp in length
*   70695    70794: gap of unknown length
*   70795    73386: contig of 2592 bp in length
*   73387    73486: gap of unknown length
*   73487    75692: contig of 2206 bp in length
*   75693    75792: gap of unknown length
*   75793    77929: contig of 2137 bp in length
*   77930    78029: gap of unknown length
*   78030    79296: contig of 1267 bp in length
*   79297    79396: gap of unknown length
*   79397    82238: contig of 2842 bp in length
*   82239    82338: gap of unknown length
*   82339    84658: contig of 2320 bp in length
*   84659    84758: gap of unknown length
*   84759    87336: contig of 2578 bp in length
*   87337    87436: gap of unknown length
*   87437    89496: contig of 2060 bp in length
*   89497    89596: gap of unknown length
*   89597    90917: contig of 1321 bp in length
*   90918    91017: gap of unknown length
*   91018    92625: contig of 1608 bp in length
*   92626    92725: gap of unknown length
*   92726    93955: contig of 1230 bp in length
*   93956    94055: gap of unknown length
*   94056    95238: contig of 1183 bp in length
*   95239    95338: gap of unknown length
*   95339    96506: contig of 1168 bp in length
*   96507    96606: gap of unknown length
*   96607    98414: contig of 1808 bp in length
*   98415    98514: gap of unknown length
*   98515   100270: contig of 1756 bp in length
*  100271   100370: gap of unknown length
*  100371   101914: contig of 1544 bp in length
*  101915   102014: gap of unknown length
*  102015   103378: contig of 1364 bp in length
*  103379   103478: gap of unknown length
*  103479   104660: contig of 1182 bp in length
*  104661   104760: gap of unknown length
*  104761   106384: contig of 1624 bp in length
*  106385   106484: gap of unknown length
*  106485   107638: contig of 1154 bp in length
*  107639   107738: gap of unknown length
*  107739   109065: contig of 1327 bp in length
*  109066   109165: gap of unknown length
*  109166   110732: contig of 1567 bp in length
*  110733   110832: gap of unknown length
*  110833   112101: contig of 1269 bp in length
*  112102   112201: gap of unknown length
*  112202   113509: contig of 1308 bp in length
*  113510   113609: gap of unknown length
*  113610   115100: contig of 1491 bp in length
*  115101   115200: gap of unknown length
*  115201   116535: contig of 1335 bp in length
*  116536   116635: gap of unknown length
*  116636   118099: contig of 1464 bp in length
*  118100   118199: gap of unknown length
*  118200   119225: contig of 1026 bp in length
*  119226   119325: gap of unknown length
*  119326   121099: contig of 1774 bp in length
*  121100   121199: gap of unknown length
*  121200   122556: contig of 1357 bp in length
*  122557   122656: gap of unknown length
*  122657   124309: contig of 1653 bp in length
*  124310   124409: gap of unknown length
*  124410   125647: contig of 1238 bp in length
*  125648   125747: gap of unknown length
*  125748   126786: contig of 1039 bp in length
*  126787   126886: gap of unknown length
*  126887   128160: contig of 1274 bp in length
*  128161   128260: gap of unknown length
*  128261   129777: contig of 1517 bp in length
*  129778   129877: gap of unknown length
*  129878   130976: contig of 1099 bp in length
*  130977   131076: gap of unknown length
*  131077   132445: contig of 1369 bp in length
*  132446   132545: gap of unknown length
*  132546   133844: contig of 1299 bp in length
*  133845   133944: gap of unknown length
*  133945   135012: contig of 1068 bp in length

  Query Match             4.7%;  Score 23;  DB 2;  Length 138556;
  Best Local Similarity   100.0%;  Pred. No. 0.22;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    326 aagggcagaagcaaagccttctc 348
          |||||||||||||||||||||||
Db  39119 AAGGGCAGAAGCAAAGCCTTCTC 39097


Search completed: June 27, 2002, 13:08:47
Job time: 6695 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 12:06:47 ; Search time 204.48 Seconds
(without alignments)
4089.091 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	487	100.0	487	22	AAH84016	Human olfactory re
2	486	99.8	486	22	AAH32285	Human olfactory re
3	236	48.5	485	22	AAH83986	Pan troglodytes ol
4	185	38.0	487	22	AAH83992	Hylobates lar olfa
5	29	6.0	988	22	AAH32168	Human olfactory re
6	22	4.5	953	22	AAI99556	Human expressed po
7	19	3.9	915	22	AAH32012	Human olfactory re
8	19	3.9	915	22	AAH32436	Human olfactory re
9	19	3.9	918	22	AAS42358	Human cDNA encodin
10	19	3.9	1037	11	AAQ04705	USP-Promoter-casse
11	19	3.9	1160	11	AAQ04706	USP-signalpeptide
12	19	3.9	1166	11	AAQ04704	USP-Promoter-casse
c 13	19	3.9	4809	23	ABL14992	Drosophila melanog
14	18	3.7	494	20	AAZ24554	Human lung tumor a
15	18	3.7	494	21	AAC65793	Human lung cancer-
16	18	3.7	714	22	AAS35473	Human cardiovascul
17	18	3.7	778	22	AAS36434	Human cardiovascul
18	18	3.7	921	22	AAH31745	Human olfactory re
19	18	3.7	924	22	AAS42225	Human cDNA encodin
20	18	3.7	924	22	AAS42336	Human cDNA encodin
21	18	3.7	967	22	AAH31747	Human olfactory re
22	18	3.7	980	22	AAD13713	Human G-protein co
23	18	3.7	980	22	AAD13714	Human G-protein co
c 24	18	3.7	2411	23	ABL19014	Drosophila melanog
25	18	3.7	2635	18	AAT83966	DNA encoding a mtl
c 26	18	3.7	5190	23	ABL05582	Drosophila melanog
c 27	18	3.7	5277	23	ABL05572	Drosophila melanog
28	18	3.7	5864	24	ABL33441	Human immune syste
29	18	3.7	6060	18	AAV74845	Staphylococcus aur
c 30	18	3.7	8861	22	AAS27785	DNA encoding novel
c 31	18	3.7	580073	18	AAT58840	Mycoplasma genital
32	17	3.5	119	21	AAC24832	Human secreted pro
33	17	3.5	300	21	AAA00216	Human colon cancer
c 34	17	3.5	315	22	AAK73673	Human immune/haema
c 35	17	3.5	315	22	AAK73674	Human immune/haema
c 36	17	3.5	317	22	AAK58786	Human immune/haema
c 37	17	3.5	400	18	AAV78417	Staphylococcus aur
c 38	17	3.5	420	20	AAX85057	Human secreted pro
c 39	17	3.5	476	22	AAS43417	Human G-Protein Co
c 40	17	3.5	482	22	ABA75636	Human foetal liver
c 41	17	3.5	482	22	ABA40244	Probe #18710 for g
c 42	17	3.5	482	22	AAK24238	Human brain expres
c 43	17	3.5	482	22	AAK50274	Human bone marrow
c 44	17	3.5	482	22	AAI27348	Probe #17281 for g
c 45	17	3.5	482	22	AAI56232	Probe #24918 used

ALIGNMENTS

RESULT 1
AAH84016
ID AAH84016 standard; DNA; 487 BP.
XX
AC AAH84016;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human olfactory receptor encoding gene 4.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance; ds.
XX
OS Homo sapiens.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB02017.
XX
PR 22-DEC-1999; 99US-0171746.
PR 21-DEC-2000; 2000US-0747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR P-PSDB; AAG98521.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful

PT for analysis odours e.g. in food processing and perfumery -
XX
PS Claim 1; Page 273-274; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 487 BP; 131 A; 105 C; 77 G; 174 T; 0 other;

Query Match 100.0%; Score 487; DB 22; Length 487;
Best Local Similarity 100.0%; Pred. No. 8.2e-234;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240

Qy    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 tgcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420

Qy    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 atctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattcc 480

Qy    481 cctgcta 487
          |||||||
Db    481 cctgcta 487
```

RESULT 2
AAH32285
ID AAH32285 standard; DNA; 486 BP.
XX
AC AAH32285;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 858.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 523; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 486 BP; 131 A; 105 C; 77 G; 173 T; 0 other;

Query Match 99.8%; Score 486; DB 22; Length 486;
Best Local Similarity 100.0%; Pred. No. 2.6e-233;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctcag 60

Qy     62 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttacta 120

Qy    122 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttacaa 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ttgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttacaa 180

Qy    182 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ctgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttggt 240

Qy    242 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gcttttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctcttt 300

Qy    302 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgcc 360

Qy    362 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 421
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 catctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgca 420

Qy    422 tctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattccc 481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 tctggcttagctgaagaccaagacaaagtgtattctctgttttacacgattataattccc 480
```

```
Qy    482 ctgcta 487
          ||||||
Db    481 ctgcta 486


RESULT   3
AAH83986
ID   AAH83986 standard; DNA; 485 BP.
XX
AC   AAH83986;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Pan troglodytes olfactory receptor encoding gene 10.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance; ds.
XX
OS   Pan troglodytes.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   P-PSDB; AAG98501.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 1; Page 227; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence 485 BP; 130 A; 107 C; 77 G; 171 T; 0 other;

  Query Match             48.5%;  Score 236;  DB 22;  Length 485;
  Best Local Similarity   99.1%;  Pred. No. 3.5e-108;
  Matches  436;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gttgctaagcatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaatttttaca 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaatgatatttatttttgg 240

Qy    241 tgctttatacaaatacccactttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    241 tgctttatacaaatacccactttaatgacgatcataatctcttatactcgtgtgctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcagcgc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420

Qy    421 atctggcttagctgaagacc 440
          ||||||||||||||||||||
Db    421 atctggcttagctgaagacc 440


RESULT   4
AAH83992
ID   AAH83992 standard; DNA; 487 BP.
XX
AC   AAH83992;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Hylobates lar olfactory receptor encoding gene 4.
XX
KW   Olfactory receptor; primate; mouse; human; food processing industry;
KW   aromas; perfumery; toxic substance; ds.
XX
OS   Hylobates lar.
XX
PN   WO200146262-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-IB02017.
XX
PR   22-DEC-1999;   99US-0171746.
PR   21-DEC-2000; 2000US-0747155.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Rouquier S,  Giorgi D;
XX
DR   WPI; 2001-381911/40.
DR   P-PSDB; AAG98506.
XX
PT   Nucleic acids encoding primate and murine olfactory receptors, useful
PT   for analysis odours e.g. in food processing and perfumery -
XX
PS   Claim 1; Page 238; 482pp; English.
XX
CC   The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC   genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC   primate species, mouse and human. The nucleic acids and receptors may be
CC   used in the food processing industry (e.g. for the detection of aromas,
CC   quality control and sample analysis), in perfumery (e.g. for the analysis
CC   or comparison of perfumes) and in the environment (e.g. for the detection
CC   of toxic substances and/or trapping of odours).
XX
SQ   Sequence 487 BP; 129 A; 107 C; 78 G; 173 T; 0 other;

  Query Match             38.0%;  Score 185;  DB 22;  Length 487;
  Best Local Similarity   98.9%;  Pred. No. 1.2e-82;
  Matches  435;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tgtagccatatgtaatcccttgctttatccagtgatgatgtccaacaaactcagcgctca 60

Qy     61 gttgctaagtatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120
```

```
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gttgctaagcatttcatatgtaattggtttcctgcatcctctggttcatgtgagtttact 120

Qy    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 attgcgactaactttctgcaggtttaacataatacattatttctactgtgaaattttaca 180

Qy    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaataatatttatttttgg 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    181 actgttcaaaatttcatgcaatggtccatctattaacgcactaatgatatttatttttgg 240

Qy    241 tgcttttatacaaatacccacttttaatgactatcataatctcttatactcgtgtgctctt 300
          |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db    241 tgcttttatacaaatacccacttttaatgacgatcataatctcttattctcgtgtgctctt 300

Qy    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcggcgc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    301 tgatattctgaaaaaaaagtctgaaaagggcagaagcaaagccttctccacatgcagcgc 360

Qy    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccatctgctttctgtctcattgtactacggaactctgatcttcatgtatgtgcgtcctgc 420

Qy    421 atctggcttagctgaagacc 440
          ||||||||||||||||||||
Db    421 atctggcttagctgaagacc 440


RESULT   5
AAH32168
ID   AAH32168 standard; DNA; 988 BP.
XX
AC   AAH32168;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 741.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 8; Page 478; 1857pp; English.
XX
CC   The present sequence is one of a number of isolated polynucleotides
CC   which encode polypeptides involved in olfactory sensation. The
CC   polynucleotides can be used in screening for olfactory agonists and
CC   antagonists. The methods allow for the determination of primary
CC   scents and the identification of the odour receptors used to detect
CC   these primary scents. The methods also enable determination of
CC   secondary scents and the identification of combinations of odour
CC   receptors that are involved in detecting such secondary scents.
CC   This enables the construction of a scent representation (also called
CC   a scent fingerprint or scent profile), which may be used to re-create
CC   and edit scents. Libraries of olfactory receptors are useful for
CC   determining the interaction pattern of a composition with the receptors,
CC   and can be used for determining differences in the olfactory faculties
CC   of different individuals.
XX
SQ   Sequence 988 BP; 250 A; 231 C; 165 G; 342 T; 0 other;


  Query Match             6.0%;  Score 29;  DB 22;  Length 988;
  Best Local Similarity   100.0%;  Pred. No. 0.00014;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    152 atacattatttctactgtgaaattttaca 180
          ||||||||||||||||||||||||||||||
Db    521 atacattatttctactgtgaaattttaca 549


RESULT   6
AAI99556
ID   AAI99556 standard; cDNA; 953 BP.
XX
AC   AAI99556;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polynucleotide SEQ ID NO 19.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200155387-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01310.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
```

PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR P-PSDB; AAM99944.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/ or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 1; SEQ ID NO 19; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 953 BP; 242 A; 191 C; 181 G; 339 T; 0 other;

Query Match 4.5%; Score 22; DB 22; Length 953;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    150 taatacattatttctactgtga 171
          |||||||||||||||||||||
Db    496 taatacattatttctactgtga 517
```

RESULT 7
AAH32012
ID AAH32012 standard; DNA; 915 BP.
XX
AC AAH32012;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 585.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 414-415; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 915 BP; 193 A; 253 C; 187 G; 282 T; 0 other;

Query Match 3.9%; Score 19; DB 22; Length 915;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    337 caaagccttctccacatgc 355
          |||||||||||||||||||
Db    702 caaagccttctccacatgc 720
```

RESULT 8
AAH32436
ID AAH32436 standard; DNA; 915 BP.
XX
AC AAH32436;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 1009.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 583; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 915 BP; 194 A; 252 C; 186 G; 283 T; 0 other;

Query Match 3.9%; Score 19; DB 22; Length 915;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    337 caaagccttctccacatgc 355
          |||||||||||||||||||
Db    702 caaagccttctccacatgc 720
```

RESULT 9
AAS42358
ID AAS42358 standard; cDNA; 918 BP.
XX
AC AAS42358;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human cDNA encoding olfactory receptor AOLFR163.

XX
KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS Homo sapiens.
XX
PN WO200168805-A2.
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07771.
XX
PR 13-MAR-2000; 2000US-0188914.
PR 24-MAR-2000; 2000US-0192033.
PR 12-APR-2000; 2000US-0198474.
PR 24-APR-2000; 2000US-0199335.
PR 26-MAY-2000; 2000US-0207702.
PR 23-JUN-2000; 2000US-0213849.
PR 16-AUG-2000; 2000US-0226534.
PR 07-SEP-2000; 2000US-0230732.
PR 07-FEB-2001; 2001US-0266862.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR P-PSDB; AAU24665.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors,
PT useful for screening for compounds involved in olfactory sensation,
PT where the compounds can be used in the food, pharmaceutical and
PT cosmetic industries to customise odours -
XX
PS Claim 1; Page 150-151; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence encodes a human olfactory receptor of
CC the invention.
XX
SQ Sequence 918 BP; 196 A; 253 C; 185 G; 284 T; 0 other;

Query Match 3.9%; Score 19; DB 22; Length 918;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 337 caaagccttctccacatgc 355
||||||||||||||||||||
Db 702 caaagccttctccacatgc 720

RESULT 10
AAQ04705
ID AAQ04705 standard; DNA; 1037 BP.
XX
AC AAQ04705;
XX
DT 12-OCT-1990 (first entry)
XX
DE USP-Promoter-cassette USP-Pr.T7.1.
XX
KW Foreign DNA incorporation; recombinant DNA techniques;
KW higher plant genome; legumin; USP-Pr.T7-1; ss.
XX
PN DE3920034-A.
XX
PD 31-MAY-1990.
XX
PF 20-JUN-1989; 89DE-3920034.
XX
PR 19-SEP-1988; 88DD-0319887.
XX
PA (PFLA-) VE KOMB PFLANZENZUC.
XX
PI Bassuner R, Baumlein H, Muntz K, Hai NV, Wobus U;
XX
DR WPI; 1990-172459/23.
XX
PT Incorporation of DNA into higher plant genome - by specified
PT recombinant DNA techniques.
XX
PS Disclosure; ; pp; German.
XX
CC The unique BglII-Ort (720-725) site is for ligating foreign DNA and the
CC HindIII-Ort in the 3' polylinker (1032-1037) for cloning the
CC cassette in the Ti-vector pGA471. Agrobacterium tumefaciens is
CC transfected.
CC See also AAQ04703-Q04706.
XX
SQ Sequence 1037 BP; 338 A; 174 C; 166 G; 359 T; 0 other;

Query Match 3.9%; Score 19; DB 11; Length 1037;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 460 gttttacacgattataatt 478
||||||||||||||||||||
Db 687 gttttacacgattataatt 705

RESULT 11
AAQ04706
ID AAQ04706 standard; DNA; 1160 BP.
XX
AC AAQ04706;
XX
DT 12-OCT-1990 (first entry)
XX
DE USP-signalpeptide cassette USP-Sig.T7.
XX
KW Foreign DNA incorporation; recombinant DNA techniques;
KW higher plant genome; signalpeptide; USP-Sig.T7.; ss.
XX
FH Key Location/Qualifiers
FT CDS 708..877
FT /*tag= a
FT /product=signalpeptide
FT intron 747..817
XX
PN DE3920034-A.
XX
PD 31-MAY-1990.
XX
PF 20-JUN-1989; 89DE-3920034.
XX
PR 19-SEP-1988; 88DD-0319887.
XX
PA (PFLA-) VE KOMB PFLANZENZUC.
XX
PI Bassuner R, Baumlein H, Muntz K, Hai NV, Wobus U;
XX
DR WPI; 1990-172459/23.
DR P-PSDB; AAR05199.
XX
PT Incorporation of DNA into higher plant genome - by specified
PT recombinant DNA techniques.

```
XX
PS   Disclosure; ; pp; German.
XX
CC   The unique BglII-Ort (890-895) site is for
CC   ligating foreign DNA and the HindIII-Ort in the 3' polylinker
CC   (1155-1160) for cloning the cassette in the Ti-vector pGA471.
CC   The cassette is cloned into the binary Ti-vectors pGA471 and
CC   Agrobacterium tumefaciens is transfected.
CC   See also AAQ04703-Q04706.
XX
SQ   Sequence 1160 BP; 369 A; 192 C; 188 G; 411 T; 0 other;


  Query Match             3.9%;  Score 19;  DB 11;  Length 1160;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    460 gttttacacgattataatt 478
          ||||||||||||||||||
Db    672 gttttacacgattataatt 690


RESULT  12
AAQ04704
ID   AAQ04704 standard; DNA; 1166 BP.
XX
AC   AAQ04704;
XX
DT   12-OCT-1990  (first entry)
XX
DE   USP-Promoter-cassette USP-Pr.T7.2.
XX
KW   Foreign DNA incorporation; recombinant DNA techniques;
KW   higher plant genome; legumin; USP-Pr.T7-2; ss.
XX
PN   DE3920034-A.
XX
PD   31-MAY-1990.
XX
PF   20-JUN-1989;   89DE-3920034.
XX
PR   19-SEP-1988;   88DD-0319887.
XX
PA   (PFLA-) VE KOMB PFLANZENZUC.
XX
PI   Bassuner R,  Baumlein H,  Muntz K,  Hai NV,  Wobus U;
XX
DR   WPI; 1990-172459/23.
XX
PT   Incorporation of DNA into higher plant genome - by specified
PT   recombinant DNA techniques.
XX
PS   Disclosure; ; pp; German.
XX
CC   The unique BglII-Ort (720-725) site is for ligating foreign DNA and the
CC   HindIII-Ort in the 3' polylinker  (1261-1266) for cloning the
CC   cassette in the Ti-vector pGA471. Agrobacterium tumefaciens is
CC   transfected.
CC   See also AAQ04703-Q04706.
XX
SQ   Sequence 1166 BP; 381 A; 208 C; 196 G; 381 T; 0 other;


  Query Match             3.9%;  Score 19;  DB 11;  Length 1166;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    460 gttttacacgattataatt 478
          ||||||||||||||||||
Db    687 gttttacacgattataatt 705


RESULT  13
ABL14992/c
ID   ABL14992 standard; cDNA; 4809 BP.
XX
AC   ABL14992;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 39458.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB70889.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 39458; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 4809 BP; 1309 A; 909 C; 1025 G; 1566 T; 0 other;


  Query Match             3.9%;  Score 19;  DB 23;  Length 4809;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    219 cactaataatatttatttt 237
          ||||||||||||||||||
Db    470 CACTAATAATATTTATTTT 452


RESULT  14
AAZ24554
ID   AAZ24554 standard; cDNA; 494 BP.
XX
AC   AAZ24554;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human lung tumor associated polynucleotide.
XX
KW   Human; lung tumor; lung cancer; T cell stimulation; ss.
XX
OS   Homo sapiens.
```

XX
PN WO9947674-A2.
XX
PD 23-SEP-1999.
XX
PF 17-MAR-1999; 99WO-US05798.
XX
PR 18-MAR-1998; 98US-0040802.
PR 18-MAR-1998; 98US-0040984.
PR 27-JUL-1998; 98US-0123912.
PR 27-JUL-1998; 98US-0123933.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Wang T;
XX
DR WPI; 1999-571839/48.
XX
PT New isolated lung tumor polynucleotides, used to develop products for
PT the treatment, prevention and monitoring the progression of lung cancer
PT -
XX
PS Claim 1; Page 81; 148pp; English.
XX
CC The invention provides isolated human lung tumor nucleic acids and
CC polypeptides. The polypeptides can be used for the treatment of lung
CC cancer. The polypeptides and polynucleotides can be used to stimulate T
CC cells or antigen presenting cells for use in the treatment of lung
CC cancer. The polypeptides and monoclonal antibodies specific for the
CC polypeptides can also be used to inhibit the development of lung cancer.
CC Agents which bind the polypeptides can be used for detecting lung cancer
CC and for monitoring the progression of lung cancer.
XX
SQ Sequence 494 BP; 164 A; 75 C; 81 G; 171 T; 3 other;

Query Match 3.7%; Score 18; DB 20; Length 494;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 tgaaaagggcagaagcaa 339
||||||||||||||||||
Db 166 tgaaaagggcagaagcaa 183

RESULT 15
AAC65793
ID AAC65793 standard; cDNA; 494 BP.
XX
AC AAC65793;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human lung cancer-associated cDNA LST-S1-H2-1G.
XX
KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection; ss.
XX
OS Homo sapiens.
XX
PN WO200061612-A2.
XX
PD 19-OCT-2000.
XX
PF 03-APR-2000; 2000WO-US08896.
XX
PR 02-APR-1999; 99US-0285479.
PR 17-DEC-1999; 99US-0466396.
PR 30-DEC-1999; 99US-0476496.
PR 10-JAN-2000; 2000US-0480884.
PR 22-FEB-2000; 2000US-0510376.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L;
XX
DR WPI; 2000-628399/60.
XX
PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT protein is used for detecting and monitoring progression of lung cancer
PT in a patient -
XX
PS Claim 1a; Page 114; 261pp; English.
XX
CC This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer.
XX
SQ Sequence 494 BP; 164 A; 75 C; 81 G; 171 T; 3 other;

Query Match 3.7%; Score 18; DB 21; Length 494;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 tgaaaagggcagaagcaa 339
||||||||||||||||||
Db 166 tgaaaagggcagaagcaa 183

Search completed: June 27, 2002, 13:12:29
Job time: 3942 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 27, 2002, 11:20:57 ; Search time 43.38 Seconds
(without alignments)
2757.573 Million cell updates/sec

Title: US-09-747-155-224
Perfect score: 487
Sequence: 1 tgtagccatatgtaatccct.....cgattataattcccctgcta 487

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18	3.7	494	4	US-09-040-984-54	Sequence 54, Appl
	2	18	3.7	494	4	US-09-123-912-54	Sequence 54, Appl
c	3	17	3.5	4247	1	US-08-061-465-2	Sequence 2, Appli
c	4	17	3.5	6828	1	US-08-061-465-1	Sequence 1, Appli
c	5	16	3.3	546	4	US-08-971-090-1	Sequence 1, Appli
c	6	16	3.3	546	4	US-09-046-479-3	Sequence 3, Appli
	7	16	3.3	600	5	PCT-US93-10418-1	Sequence 1, Appli
c	8	16	3.3	1128	1	US-08-448-744-6	Sequence 6, Appli
c	9	16	3.3	1992	3	US-08-776-265-6	Sequence 6, Appli
c	10	16	3.3	2481	1	US-08-467-568-1	Sequence 1, Appli
c	11	16	3.3	2481	2	US-09-030-582-1	Sequence 1, Appli
c	12	16	3.3	2481	5	PCT-US94-09051-1	Sequence 1, Appli
c	13	16	3.3	3100	1	US-08-296-362-1	Sequence 1, Appli
c	14	16	3.3	4565	3	US-08-776-265-1	Sequence 1, Appli
c	15	16	3.3	5024	1	US-08-920-812-7	Sequence 7, Appli
c	16	16	3.3	5024	1	US-08-920-827-7	Sequence 7, Appli
c	17	16	3.3	5024	1	US-08-921-177-7	Sequence 7, Appli
c	18	16	3.3	5024	1	US-08-362-577C-7	Sequence 7, Appli
c	19	16	3.3	5024	2	US-08-920-828-7	Sequence 7, Appli
c	20	15	3.1	244	1	US-08-222-177A-29	Sequence 29, Appl
	21	15	3.1	471	2	US-08-502-535B-1	Sequence 1, Appli
	22	15	3.1	471	2	US-08-908-005A-1	Sequence 1, Appli
	23	15	3.1	471	4	US-08-558-818-6	Sequence 6, Appli
	24	15	3.1	471	4	US-08-974-469A-6	Sequence 6, Appli
	25	15	3.1	471	4	US-08-832-180-3	Sequence 3, Appli
	26	15	3.1	471	4	US-08-832-198-10	Sequence 10, Appl
	27	15	3.1	471	4	US-09-253-523-1	Sequence 1, Appli
	28	15	3.1	471	4	US-09-251-911-1	Sequence 1, Appli
	29	15	3.1	510	2	US-08-820-170A-23	Sequence 23, Appl
	30	15	3.1	510	3	US-09-055-699-23	Sequence 23, Appl
	31	15	3.1	510	4	US-09-273-565-23	Sequence 23, Appl
	32	15	3.1	510	4	US-09-565-538-23	Sequence 23, Appl
	33	15	3.1	617	2	US-08-820-170A-24	Sequence 24, Appl
	34	15	3.1	617	3	US-09-055-699-24	Sequence 24, Appl
	35	15	3.1	617	4	US-09-273-565-24	Sequence 24, Appl
	36	15	3.1	617	4	US-09-565-538-24	Sequence 24, Appl
	37	15	3.1	662	4	US-08-998-416-185	Sequence 185, App
c	38	15	3.1	663	4	US-08-998-416-191	Sequence 191, App
	39	15	3.1	665	4	US-08-998-416-937	Sequence 937, App
c	40	15	3.1	669	2	US-08-896-410-3	Sequence 3, Appli
	41	15	3.1	701	4	US-08-998-416-701	Sequence 701, App
	42	15	3.1	711	4	US-08-998-416-786	Sequence 786, App
	43	15	3.1	724	4	US-08-998-416-683	Sequence 683, App
	44	15	3.1	732	4	US-08-998-416-1036	Sequence 1036, Ap
	45	15	3.1	767	4	US-08-998-416-472	Sequence 472, App

ALIGNMENTS

```
RESULT 1
US-09-040-984-54
; Sequence 54, Application US/09040984
; Patent No. 6210883
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Wang, TongTong
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS
;    TITLE OF INVENTION:  OF LUNG CANCER
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/040,984
;      FILING DATE:  18-MAR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.456
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-622-4900
;      TELEFAX:  206-282-6031
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-040-984-54


  Query Match             3.7%;  Score 18;  DB 4;  Length 494;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     322 tgaaaagggcagaagcaa 339
           ||||||||||||||||||
Db     166 TGAAAAGGGCAGAAGCAA 183
```

```
RESULT   2
US-09-123-912-54
; Sequence 54, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE:  210121.455C1
;  CURRENT APPLICATION NUMBER: US/09/123,912A
;  CURRENT FILING DATE:  1998-07-27
;  PRIOR APPLICATION NUMBER:  09/040,802
;  PRIOR FILING DATE:  1998-03-18
;  NUMBER OF SEQ ID NOS:  114
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 494
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (431)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (442)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (445)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-54


  Query Match             3.7%;  Score 18;  DB 4;  Length 494;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     322 tgaaaagggcagaagcaa 339
           ||||||||||||||||||
Db     166 tgaaaagggcagaagcaa 183


RESULT   3
US-08-061-465-2/c
; Sequence 2, Application US/08061465
; Patent No. 5625049
;  GENERAL INFORMATION:
;    APPLICANT:  Monroe, Stephan S.
;    APPLICANT:  Glass, Roger I.
;    APPLICANT:  Koopmans, Marion
;    APPLICANT:  Jiang, Baoming
;    TITLE OF INVENTION:  NUCLEIC ACIDS ENCODING HUMAN ASTROVIRUS
;    TITLE OF INVENTION:  SEROTYPE II AND USES THEREOF
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  127 Peachtree Street, N.E.
;      STREET:  Suite 1200
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/061,465
;      FILING DATE:  19930512
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perryman, David G.
;      REGISTRATION NUMBER:  33,438
;      REFERENCE/DOCKET NUMBER:  1414.069
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4247 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Human Astrovirus
;      STRAIN:  Serotype 2
US-08-061-465-2


  Query Match             3.5%;  Score 17;  DB 1;  Length 4247;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     302 gatattctgaaaaaaaa 318
           |||||||||||||||||
Db    3021 GATATTCTGAAAAAAAA 3005


RESULT   4
US-08-061-465-1/c
; Sequence 1, Application US/08061465
; Patent No. 5625049
;  GENERAL INFORMATION:
;    APPLICANT:  Monroe, Stephan S.
;    APPLICANT:  Glass, Roger I.
;    APPLICANT:  Koopmans, Marion
;    APPLICANT:  Jiang, Baoming
;    TITLE OF INVENTION:  NUCLEIC ACIDS ENCODING HUMAN ASTROVIRUS
;    TITLE OF INVENTION:  SEROTYPE II AND USES THEREOF
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  127 Peachtree Street, N.E.
;      STREET:  Suite 1200
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/061,465
;      FILING DATE:  19930512
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perryman, David G.
;      REGISTRATION NUMBER:  33,438
;      REFERENCE/DOCKET NUMBER:  1414.069
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6828 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
```

```
;    ORIGINAL SOURCE:
;      ORGANISM:  Human Astrovirus
;      STRAIN:  Serotype 2
;    POSITION IN GENOME:
;      UNITS:  100%
US-08-061-465-1


  Query Match             3.5%;  Score 17;  DB 1;  Length 6828;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    302 gatattctgaaaaaaaa 318
          ||||||||||||||||||
Db   3103 GATATTCTGAAAAAAAA 3087


RESULT   5
US-08-971-090-1/c
; Sequence 1, Application US/08971090
; Patent No. 6228579
; GENERAL INFORMATION:
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Forsyth, R. Allyn
;  TITLE OF INVENTION: METHOD FOR IDENTIFYING MICROBIAL PROLIFERATION GENES
;  FILE REFERENCE: 07252/008001
;  CURRENT APPLICATION NUMBER: US/08/971,090
;  CURRENT FILING DATE: 1997-11-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 546
;   TYPE: RNA
;   ORGANISM: E. coli
US-08-971-090-1


  Query Match             3.3%;  Score 16;  DB 4;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     20 ttgctttatccagtga 35
          ||||||||||||||||
Db    268 TTGCTTTATCCAGTGA 253


RESULT   6
US-09-046-479-3/c
; Sequence 3, Application US/09046479
; Patent No. 6291653
;  GENERAL INFORMATION:
;    APPLICANT:  Sheppard, Paul O.
;    APPLICANT:  Deisher, Theresa A.
;    TITLE OF INVENTION:  MOTILIN HOMOLOGS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/046,479
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sawislak, Deborah A
;      REGISTRATION NUMBER:  37,438
;      REFERENCE/DOCKET NUMBER:  97-04
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6672
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  546 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  40...396
;      OTHER INFORMATION:
US-09-046-479-3


  Query Match             3.3%;  Score 16;  DB 4;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     88 tttcctgcatcctctg 103
          ||||||||||||||||
Db    160 TTTCCTGCATCCTCTG 145


RESULT   7
PCT-US93-10418-1
; Sequence 1, Application PC/TUS9310418
;  GENERAL INFORMATION:
;    APPLICANT:  Ziegler, Steven F.
;    APPLICANT:  Hjerrild, Kathryn A.
;    TITLE OF INVENTION:  Activation Antigen CD69
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/10418
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seese, Kathryn A.
;      REGISTRATION NUMBER:  32,172
;      REFERENCE/DOCKET NUMBER:  2610-WO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 233-0644
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  600 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
```

```
;       NAME/KEY:  CDS
;       LOCATION:  1..600
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  1..597
PCT-US93-10418-1


  Query Match             3.3%;  Score 16;  DB 5;  Length 600;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    157 ttatttctactgtgaa 172
          ||||||||||||||||
Db    293 TTATTTCTACTGTGAA 308


RESULT   8
US-08-448-744-6/c
; Sequence 6, Application US/08448744
; Patent No. 5616484
;  GENERAL INFORMATION:
;    APPLICANT:  XU, Shuang-yong
;    TITLE OF INVENTION:  Cloning And Expression of The ApaLI
;    TITLE OF INVENTION:  Restriction Endonuclease
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  New England Biolabs, Inc.
;      STREET:  32 Tozer Road
;      CITY:  Beverly
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  01915
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,744
;      FILING DATE:  24-May-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Gregory D.
;      REGISTRATION NUMBER:  30901
;      REFERENCE/DOCKET NUMBER:  NEB-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (508) 927-5054
;      TELEFAX:  (508) 927-1705
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-448-744-6


  Query Match             3.3%;  Score 16;  DB 1;  Length 1128;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    391 aactctgatcttcatg 406
          ||||||||||||||||
Db   1096 AACTCTGATCTTCATG 1081


RESULT   9
US-08-776-265-6/c
; Sequence 6, Application US/08776265
; Patent No. 6001631
;  GENERAL INFORMATION:
;    APPLICANT:  BLANCHE, Francis
;    APPLICANT:  CAMERON, Beatrice
;    APPLICANT:  CROUZET, Joel
;    APPLICANT:  FAMECHON, Alain
;    APPLICANT:  FERRERO, Lucia
;    TITLE OF INVENTION:  No. 6001631el Topoisomerase IV, Corresponding
;    TITLE OF INVENTION:  Nucleotide Sequences and Uses Thereof
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner, L.L.P.
;      STREET:  1300 I. Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/776,265
;      FILING DATE:  24-JAN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Einaudi, Carol P.
;      REGISTRATION NUMBER:  32,220
;      REFERENCE/DOCKET NUMBER:  03806.0394-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 408-4000
;      TELEFAX:  (202) 408-4444
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1992 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-776-265-6


  Query Match             3.3%;  Score 16;  DB 3;  Length 1992;
  Best Local Similarity   100.0%;  Pred. No. 53;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    463 ttacacgattataatt 478
          ||||||||||||||||
Db   1495 TTACACGATTATAATT 1480


RESULT  10
US-08-467-568-1/c
; Sequence 1, Application US/08467568
; Patent No. 5817477
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    TITLE OF INVENTION:  ADRENERGIC RECEPTOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan, Cecchi,
;      ADDRESSEE:  Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07068-1739
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/467,568
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-324
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2481 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  101..1687
US-08-467-568-1


  Query Match             3.3%;  Score 16;  DB 1;  Length 2481;
  Best Local Similarity   100.0%;  Pred. No. 53;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 aataatatttattttt 238
          ||||||||||||||||
Db   2467 AATAATATTTATTTTT 2452


RESULT  11
US-09-030-582-1/c
; Sequence 1, Application US/09030582
; Patent No. 5994506
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    TITLE OF INVENTION:  ADRENERGIC RECEPTOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan, Cecchi,
;      ADDRESSEE:  Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07068-1739
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/030,582
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/467,568
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-324
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2481 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  101..1687
US-09-030-582-1


  Query Match             3.3%;  Score 16;  DB 2;  Length 2481;
  Best Local Similarity   100.0%;  Pred. No. 53;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 aataatatttattttt 238
          ||||||||||||||||
Db   2467 AATAATATTTATTTTT 2452


RESULT  12
PCT-US94-09051-1/c
; Sequence 1, Application PC/TUS9409051
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Adrenergic Receptor
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09051
;      FILING DATE:  Submitted herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2481 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
PCT-US94-09051-1


  Query Match             3.3%;  Score 16;  DB 5;  Length 2481;
  Best Local Similarity   100.0%;  Pred. No. 53;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 aataatatttattttt 238
          ||||||||||||||||
Db   2467 AATAATATTTATTTTT 2452


RESULT  13
US-08-296-362-1/c
; Sequence 1, Application US/08296362
; Patent No. 5691306
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Bergeron, John J.M.
;    APPLICANT:  Thomas, David Y.
;    APPLICANT:  Wada, Ikuo
;    TITLE OF INVENTION:  METHODS OF DETECTION AND TREATMENT OF
;    TITLE OF INVENTION:  PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
;    TITLE OF INVENTION:  PROTEIN PRODUCTION
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/296,362
;      FILING DATE:  25-AUG-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Deehr, Manya S.
;      REGISTRATION NUMBER:  37,120
;      REFERENCE/DOCKET NUMBER:  690066.401C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;      TELEX:  3723836
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3100 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  102..1883
US-08-296-362-1


  Query Match             3.3%;  Score 16;  DB 1;  Length 3100;
  Best Local Similarity   100.0%;  Pred. No. 53;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     182 ctgttcaaaatttcat 197
           ||||||||||||||||
Db    1849 CTGTTCAAAATTTCAT 1834


RESULT  14
US-08-776-265-1/c
; Sequence 1, Application US/08776265
; Patent No. 6001631
;  GENERAL INFORMATION:
;    APPLICANT:  BLANCHE, Francis
;    APPLICANT:  CAMERON, Beatrice
;    APPLICANT:  CROUZET, Joel
;    APPLICANT:  FAMECHON, Alain
;    APPLICANT:  FERRERO, Lucia
;    TITLE OF INVENTION:  No. 6001631el Topoisomerase IV, Corresponding
;    TITLE OF INVENTION:  Nucleotide Sequences and Uses Thereof
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner, L.L.P.
;      STREET:  1300 I. Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/776,265
;      FILING DATE:  24-JAN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Einaudi, Carol P.
;      REGISTRATION NUMBER:  32,220
;      REFERENCE/DOCKET NUMBER:  03806.0394-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 408-4000
;      TELEFAX:  (202) 408-4444
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4565 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-776-265-1


  Query Match             3.3%;  Score 16;  DB 3;  Length 4565;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     463 ttacacgattataatt 478
           ||||||||||||||||
Db    1535 TTACACGATTATAATT 1520


RESULT  15
US-08-920-812-7/c
; Sequence 7, Application US/08920812
; Patent No. 5763188
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,812
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
```

; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5024 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-32
US-08-920-812-7

Query Match 3.3%; Score 16; DB 1; Length 5024;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      82 aattggtttcctgcat 97
           ||||||||||||||||
Db    1574 AATTGGTTTCCTGCAT 1559
```

Search completed: June 27, 2002, 13:08:46
Job time: 6469 sec